A Review of Simulation Methods in Micro/Nanoscale Heat Conduction

Hua Bao,* Jie Chen,** Xiaokun Gu*** and Bingyang Cao****

Significant progress has been made in the past two decades about the micro/nanoscale heat conduction. Many computational methods have been developed to accommodate the needs to investigate new physical phenomena at micro/nanoscale and support the applications like microelectronics and thermoelectric materials. In this review, we first provide an introduction of state-of-the-art computational methods for micro/nanoscale conduction research. Then the physical origin of size effects in thermal transport is presented. The relationship between the different methods and their classification are discussed. In the subsequent sections, four commonly used simulation methods, including first-principles Boltzmann transport equation, molecular dynamics, non-equilibrium Green’s function, and numerical solution of phonon Boltzmann transport equation will be reviewed in details. The hybrid method and coupling scheme for multiscale heat transfer simulation are also briefly discussed.

Keywords: Heat conduction; First-principles; Boltzmann transport equation; Molecular dynamics; Non-equilibrium Green’s function

Received 4 September 2018, Accepted 7 October 2018
DOI: 10.30919/esee8c149

1. Introduction

The miniaturization of devices and structures, higher power density of the novel electronic and optic cells and more severe thermal conditions pose huge challenges to the thermal management and energy conversion issues. Specifically, heat conduction in micro/nanoscale requires sophisticated understandings and interpretations due to its distinct physical pictures from macroscopic thermal transport, and it is of great difficulty to handle multi-scale heat transfer problems using a uniform law or theory.1-2 Hence, new theories, computational methods and experimental techniques have emerged to investigate the thermal transport from nanoscale (1-100 nm) to microscale (0.1-100 μm) in the past two decades.3 Advanced experimental methods, such as optothermal Raman, electrical self-heating, T-type sensor and 3ω technique, have been developed to measure the thermal conductivity of nanowires and thin films.4-6 The measured ultrahigh thermal conductivities of graphene and carbon nanotubes (CNTs) have intrigued enormous interest in low-dimensional materials,7,8 and considerable work is devoted into laboratory studies on superlattices, nanofluids, as well as special nanostructures and interfaces at micro/nanoscale.9,10

Despite the rapid progress of experimental measurements, it is still very challenging to conduct nanoscale heat conduction experiments, and the measured values may be lack of accuracy due to many unforeseen factors.11 Therefore, computational methods are highly demanded to assist the experiment studies to explain underlying mechanisms or predict new physical phenomena that the current measurement techniques are not mature enough to observe. For instance, the thermal rectification in asymmetric graphene, which was predicted by many molecular dynamics simulations, has been realized in the experiment recently.12,13 First-principles calculations predicted the important role of low-frequency phonons in Si, which was later confirmed by transient thermal grating measurements over micron distances.14-15 Owing to the newly developed computational methods in the last twenty years, the temperature and heat flow can be resolved into phonon contributions,16-20 and the spatial and time information of phonons is more accessible to give a comprehensive understanding of ballistic, hydrodynamic, coherent, localized and other unique transport regimes.21-25 Moreover, a better knowledge of micro-nanoscale heat conduction has shed light on the promising engineering applications such as thermal diode, thermal cloak, high performance thermoelectric materials and nanocomposites.26-29

This review is focused on widely used computational methods in heat conduction at micro/nanoscale. As seen in Fig. 1, the numerical models are approximately classified into several categories based on the length and time scales of the thermal transport phenomena. Different methods are used to handle specific heat conduction problems of different range of scales, and the overlap areas of two techniques indicate that it is feasible to map
from one scale to the next scale. However, a universal equation is still lacking to deal with thermal transport issues at any scale. The methods in this review mainly involve first-principles Boltzmann transport equations (BTE), molecular dynamics (MD) simulations, non-equilibrium Green’s function (NEGF), numerical solution of phonon BTE, and hybrid methods. The classification of these methods will be discussed later and we will give a brief introduction of each approach here. The first-principles BTE methods are available for parameter-free predictions by solving the BTE with the interatomic force constants (IFCs) from density functional theory (DFT). The computational cost of the first-principles BTE is so expensive that the system size is limited to hundreds of atoms, but the advances in high performance computing have greatly accelerated the application of first-principles BTE nowadays. The predicted results from first-principles BTE are quite accurate compared to experimental data, providing guidelines for the discoveries of novel materials. MD simulations are based on the implementation of Green’s function in electron transport. The conduction at the interfaces and nanostructures, which is inspired by effective to deal with nanostructures, interfaces and other heat interactions. Despite the classical nature of MD simulations, it is accelerated the application of first-principles BTE nowadays. The predicted results from first-principles BTE are quite accurate compared to experimental data, providing guidelines for the discoveries of novel materials. MD simulations are based on the integration of the Newton’s equations for an ensemble of atoms, usually employing an empirical potential to describe atomic interactions. Despite the classical nature of MD simulations, it is effective to deal with nanostructures, interfaces and other heat transfer topics. The NEGF method is powerful to investigate heat conduction at the interfaces and nanostructures, which is inspired by the implementation of Green’s function in electron transport. The numerical solution of BTE is adopted to study mesoscopic thermal transport, and fundamental information of phonon scattering is required to incorporate into the calculations in an explicit pattern. The hybrid methods, which are not fully developed so far, aim to increase the computational efficiency of multi-scale heat transfer.

2. The origin of size effect and related simulation methods

2.1 The physical origin of size effects

At micro and nanoscale, the non-Fourier heat conduction is originated from the different physics of energy carrier transport. In most of the solid materials where micro-nanoscale conduction research is carried out, thermal energy is carried by lattice vibrations and electrons.31 Electrons are the dominating energy carriers in metallic solids while lattice vibration is the heat carrier in semiconductors and insulators. In particular, the lattice vibration in crystalline solids can be decomposed into lattice waves (normal modes) and the quantized lattice waves are treated as a quasi-particles, known as phonons.31 As shown in Fig. 2, the heat conduction in solid crystals can be understood from the transport of energy carriers (phonon and electrons). The thermal transport in solids can be understood from the simple kinetic theory for gases, from which it has been shown that the thermal conductivity can be given by the expression,31

\[
\kappa = \frac{1}{3} Cv \ell ,
\]

where \( C \) is the specific heat of particles per unit volume, \( v \) is the average particle velocity, and \( \ell \) is the mean free path (MFP), which is defined as the average travelling distance between two scattering events. The particle here can be phonons, electrons or other heat carriers.

![Fig. 2 The phonon and electron gas models of non-metals and metals. In nonmetallic crystals, heat is carried by phonon gases (gray spheres) and only phonon transport needs to be considered. In metals, both phonon and electron (brown spheres) contribute to thermal transport, and they are coupled through electron-phonon scattering.](image)

The size effect discussed below will be focusing on non-metallic crystalline solid where phonon is the major heat carrier, since it is more prominent in these solids than in metals and non-crystalline solids. Two lengths scales are very important to the micro-nanoscale size effect of heat conduction. One is the phonon wavelength and the other is the phonon MFP.3 Fig. 3 shows the accumulated thermal conductivity of silicon at three different temperatures, calculated by first-principles simulation. The horizontal axis is wavelength or MFP, and the vertical axis denotes the fraction of thermal conductivity contributed by the phonons with wavelength or MFP below this value. This figures can quantify the contribution of phonons with different wavelength or MFP to the overall thermal conductivity. It can be seen that at 300 K most of the heat is carried by phonons in the wavelength range of 0.5-10 nm, with MFP in the range of 1 nm to 10 μm. These values depend on the material type and temperature. In general, the dominating phonon wavelength and MFP are larger at lower temperatures. If the feature size of the system is much larger than the phonon MFP, then it falls into the macroscopic diffusive energy transport regime, where the heat diffusion equation works well. If the feature size is comparable
to phonon MFP, the classical size effect (phonon ballisic transport phenomenon) should be considered and the BTE can be used to describe the energy transport. If the feature size is further comparable to the wavelength, then the wave nature of phonons should be considered.

Another reason for the size effect is the interfacial thermal resistance (also known as the thermal boundary resistance or Kapitza resistance). This resistance is originated from different electronic and vibrational properties of different materials. When an energy carrier attempts to travel across the interface, it scatters at the interface. The scattering results in a temperature drop at the interface and thus gives an interfacial thermal resistance. Note that the interfacial thermal resistance occurs at atomically perfect contact and should be distinguished from the contact thermal resistance. The interfacial thermal resistance is on the order of $10^{-10}$ to $10^{-9}$ m K/W. This resistance is very small from the macroscopic point of view. However, if the interfaces are very dense, for example, when the distance between the interfaces are roughly in the range of nanometer to micron, it becomes important and even dominant.

Therefore, these new transport phenomena at small scale generally fall into the scale of nanometer to micron. This is also the length scale where we need new simulation methods to investigate the thermal transport process.

2.2 Classification of simulation methods

The simulation methods for micro/nanoscale heat conduction can be mainly categorized into two types, continuum simulations and atomistic simulations. The continuum simulations include the direct solution of BTE and the wave equation. The acoustic wave equation treats phonons as acoustic waves with linear dispersion relation and thus neglects any nonlinearity and phonon-phonon scattering. However, because phonons have broad wavelength spectrum and nonlinear dispersion relation, and because phonon-phonon scattering is important, the wave equation is rarely used to describe phonon transport except for some specific situations. Therefore it will not be discussed in this review in details. The BTE, on the other hand, is based on particle dynamics and thus neglect any wave effect. Since BTE is the governing equation for transport and scattering of energy carriers in solids (as shown in Fig. 4), it is widely used to simulate mesoscopic conduction process. BTE describes the evolution of particle distribution function $f(x, p, t)$, which denotes the fraction of particles that have position $x$ and momentum $p$ at time $t$. The general BTE has the form of

$$\frac{\partial f}{\partial t} + v \cdot \nabla f + F \cdot \nabla f = \frac{\partial}{\partial t} \left( \frac{\partial f}{\partial x} \right) + \frac{\partial}{\partial p} \left( \frac{\partial f}{\partial p} \right),$$

(2)

here $v$ denotes the velocity of particles, $F$ is the external force (e.g. gravity, external electric force, etc), $\left( \frac{\partial f}{\partial x} \right)$ is the scattering term, and $\frac{\partial f}{\partial p}$ is a source term of particles. Note that BTE is a general governing equation that applies for microscale energy carriers, such as phonons, electrons, and photons. Since phonons do not interact with external force in general, and if we further assume there is no source term, then the phonon BTE becomes,

$$\frac{\partial n}{\partial t} + v \cdot \nabla n = \frac{\partial}{\partial p} \left( \frac{\partial n}{\partial p} \right).$$

(3)

Note that here we use $n$ to denote phonon distribution function, following the convention and also distinguishing from the distribution function of other particles. To solve BTE, one needs a few input parameters, i.e., the group velocity $v$ (travelling speed) and the scattering rates $\left( \frac{\partial f}{\partial x} \right)$ of all the relevant scattering processes. The scattering is related to the nature of energy carrier, and the scattering rates can only be obtained from quantum mechanics. For phonons in non-metallic solids, the dominating scattering processes included phonon-phonon scattering, phonon-impurity scattering, and phonon-boundary scattering, as shown in Fig. 4. Note that boundary refers to the geometric boundary of the piece of crystal. The phonon-phonon scattering and phonon-impurity scattering are bulk processes, and phonon-boundary scattering is a surface or interface phenomenon.

These input parameters needed for solving BTE are related to the nature of the material (group velocity and bulk scattering terms) and thus should be taken as empirical values or more accurately obtained from atomic simulations. In addition, the boundary and interface conditions must also be provided to describe phonon-boundary scattering. Once the distribution function is obtained by solving BTE, the temperature distribution and heat flux can then be extracted. BTE is in principle not limited by the scale of simulation.
domain, as long as the scale is large enough so that each element after domain discretization contains enough particles to properly define distribution function $n(x,p,t)$. The solution of BTE converges to the macroscopic energy equation (i.e., heat diffusion equation) when the simulation domain is much larger than the transport MFP. Therefore, BTE is usually applied when the system is larger than the atomic scale and smaller than the diffusion regime.

![Fig. 4](image1.png) Different phonon scattering mechanisms. Phonon can scatter with impurities (and defects) and boundaries. Phonon can also scatter with other phonons, generally known as three-phonon scattering processes.

In contrast to the continuum simulations, atomistic simulations consider all details of atomic structures of materials. The widely used simulation methods for micro/nanoscale heat transfer include molecular dynamics simulation, lattice dynamics simulation, and non-equilibrium Green's function method. Molecular dynamics simulation tracks the evolution of atomic systems based on the Newton's second law, which is a real-space treatment of atoms. All heat conduction phenomena are naturally included in the atomic trajectories. The only problem is that it regards atoms as classical distinguishable particles, so the distribution function follows the Boltzmann distribution instead of quantum Bose-Einstein distribution. The difference between classical and quantum treatment is negligible in the high temperature regime. Lattice dynamics is a reciprocal space method which deals with normal modes (i.e., phonon modes). The phonon properties can be described by the phonon dispersion curve and phonon-phonon scattering matrix elements. The NEGF approach is an efficient tool to obtain phonon transmission coefficient when the elastic scattering is dominating in the transport process. All the three methods need the input of interatomic interaction, which can be obtained either from first-principles DFT simulations (DFT is a method to numerically solve the Schrodinger equation, or more precisely the Kohn-Sham equation) or from empirical interatomic potential. The first-principles simulation requires iterative solution of electron charge density of many-body systems and therefore usually requires very high computational cost, but it is quite accurate. It only needs the input of the initial atomic structure of the material, but the computational cost will limit such calculations to a few hundreds of atoms. The empirical potentials assume certain analytical formula for interatomic interaction, which are fitted with first-principles data or experimental results. They are much faster in calculating interatomic forces compared to first-principles simulations and can deal with millions of atoms, but the accuracy is limited. Since MD simulations of thermal transport need large atomic systems to reduce possible simulation size effect and many time steps to reach equilibrium, they are usually combined with empirical potential. Lattice dynamics and NEGF can be combined with either first-principles method or empirical potential. The first-principles method is clearly more accurate. The classification and relationship of different simulation methods for heat conduction are summarized in Fig. 5. In the subsequent sections, we will discuss these methods in details.

3. First-principles PBTE methods

Before we start to introduce the first-principles PBTE method, different terminologies used in literature should be clarified. Strictly speaking, first-principles method (also known as the ab initio method) refers to an approach to numerically solve Schrodinger equation. Anharmonic lattice dynamics or lattice dynamics takes...
the IFCs to generate phonon dispersion relation and relaxation time.\textsuperscript{45} BTE takes the phonon dispersion and relaxation time as inputs to obtain total thermal conductivity. The so-called first-principles BTE method refers to the combination of the three methods. It is sometimes called “first-principles phonon BTE method”, “first-principles anharmonic lattice dynamics method”, or even just “first-principles method” in literature. In this section, we will discuss the framework of BTE and anharmonic lattice dynamics method, which are more focused on heat transfer. The general first-principles method is well documented in a book,\textsuperscript{45} and will not be presented here. The purpose of first-principles method here is to generate IFCs. For metal calculations, it also helps to generate electronic band structures and electron-phonon matrix elements. In this section, we will first show how to obtain the relationship between thermal conductivity and phonon properties by phonon BTE. Then the lattice dynamics method will be introduced. The electron-phonon coupling is then discussed.

3.1 Thermal conductivity and BTE

The discussion will be first focused on non-metallic crystals, where phonon is the dominating heat carrier. We will show the theoretical framework of predicting lattice thermal conductivity based on anharmonic lattice dynamics and the linearized BTE.\textsuperscript{44,45} Some of our derivations in this section are adapted from Ref. 48 and 49.

Based on Fourier’s law, the thermal conductivity $\kappa$ of a material measures its ability to conduct heat and is defined by

$$ J = -\kappa \nabla T , $$

where $J$ is the heat flux vector and $\nabla T$ is the temperature gradient. The thermal conductivity $\kappa$ is a tensor in this equation. To predict thermal conductivity from atomic scale simulations, the relationship between heat flux and temperature gradient should be derived. For a phonon mode $\lambda$ that consists of both wave vector $q$ and phonon branch $n_{\lambda}$, the distribution function $f_{\lambda}$ follows the BTE, Eq. (3.).

Under linearization assumption, the diffusion term can be expressed as\textsuperscript{44,45}

$$ \nabla n_{\lambda} = \nabla \frac{T}{\beta} \frac{\partial n_{\lambda}}{\partial T} , $$

where $\nabla$ is the group velocity of phonon mode $\lambda$. The scattering term on the right hand side of Eq. (3.) also needs to be simplified. A phonon relaxation time $\tau_{\lambda}$ is defined with the following equation

$$ \left( \frac{\partial n_{\lambda}}{\partial T} \right)_{\text{stat}} = -\frac{n'_{\lambda}}{\tau_{\lambda}} , $$

where $n'_{\lambda} = n_{\lambda} - n^{0}_{\lambda}$ is the deviation of distribution function from equilibrium $n^{0}_{\lambda}$. Phonons are bosons and the equilibrium distribution follows the Bose-Einstein statistics $n^{0}_{\lambda} = \frac{1}{e^{\omega_{\lambda} / k_{B} T} - 1}$, where $h$ is reduced Planck constant and $k_{B}$ is Boltzmann constant. $\omega_{\lambda}$ denotes the phonon frequency and $T$ represents the temperature. Under the steady state and finite temperature gradient assumption, the deviation of distribution from equilibrium is small, so that $\partial n_{\lambda} / \partial T = \partial n^{0}_{\lambda} / \partial T$ in Eq. (5.).\textsuperscript{45} If we further assume steady state and neglect the time dependent term, we can get the linearized BTE,

$$ -\nabla T \frac{\partial n_{\lambda}}{\partial T} = \frac{n'_{\lambda}}{\tau_{\lambda}} . $$

In order to establish the relationship between lattice thermal conductivity and phonon properties, we further consider a finite temperature gradient established across a solid. The microscopic expression for the heat flux contributed by phonons is\textsuperscript{45}

$$ J = \frac{1}{V} \sum_{\lambda} \hbar \omega_{\lambda} v_{\lambda} n_{\lambda} , $$

where $V$ is the volume of the solid that can be calculated with $V = N \Omega$, $N_{\lambda}$ is the number of wave vector $q$ points and $\Omega$ is the volume of the unit cell. Because $n_{\lambda} = n^{0}_{\lambda} + n'_{\lambda}$ and the net heat flux under equilibrium state would be zero, Eq. (8.) can be rewritten (in $\beta$ direction) as

$$ J^{\beta} = \frac{1}{V} \sum_{\lambda} \hbar \omega_{\lambda} v_{\lambda}^{\beta} n'_{\lambda} . $$

When the temperature gradient is in $\alpha$ direction, Eq. (7.) will become $-\kappa_{\alpha} \frac{\partial T}{\partial \alpha} = \frac{n'_{\alpha}}{\tau_{\alpha}}$ and we can plug it into the Eq. (9.) to yield

$$ J^{\alpha} = \frac{1}{V} \sum_{\lambda} \hbar \omega_{\lambda} v_{\lambda}^{\alpha} v_{\lambda}^{\beta} n'_{\alpha} T_{\lambda} \frac{\partial T}{\partial \alpha} . $$

By comparing this equation with the Fourier’s law $J^{\alpha} = -\kappa^{\alpha} \frac{\partial T}{\partial \alpha}$, we can get an expression for the thermal conductivity

$$ \kappa^{\alpha} = \frac{\hbar}{V} \sum_{\lambda} \omega_{\lambda} v_{\lambda}^{\alpha} v_{\lambda}^{\beta} \tau_{\lambda} . $$

The volumetric heat capacity is related to the phonon distribution function by $c_{v} = \frac{\hbar \omega_{\lambda}}{V} \frac{\partial T}{\partial \alpha}$, and thus we can rewrite Eq. (11.) as

$$ \kappa^{\alpha} = \sum_{\lambda} c_{v} v_{\lambda}^{\alpha} v_{\lambda}^{\beta} \tau_{\lambda} . $$

Note that Eq. (12.) relates the phonon properties with the macroscopic definition of thermal conductivity, which is the more rigorous version of the kinetic theory Eq. (1.). In the following sections, we will explain how to calculate the heat capacity $c_{v}$, group velocity $v_{\lambda}$ and relaxation time $\tau_{\lambda}$ from lattice dynamics method.

3.2 Harmonic and anharmonic lattice dynamics

The harmonic lattice dynamics takes the second-order IFCs to obtain the phonon dispersion relation $\omega(q)$. If the dispersion relationship is obtained, the specific heat of mode $\lambda$ can be calculated. The group velocity of phonon mode $\lambda$ is the gradient of frequency with respect to reciprocal space coordinates $\mathbf{q}_{\lambda} = \mathbf{V}_{k} \mathbf{q}$. The phonon relaxation time should be obtained from the anharmonic lattice dynamics calculations, which need second-order (or harmonic) IFCs and higher-order (or anharmonic) IFCs. The details will be presented in this section.

For a periodic crystal, the potential energy $U$ can be expanded with atomic displacements around their equilibrium positions as the Taylor series

$$ U = U_{0} + \frac{1}{2} \sum_{\alpha,\beta} \Phi_{\alpha}^{\beta} u_{\alpha}^{\beta} + \frac{1}{3!} \sum_{\alpha,\beta,\gamma} \Psi_{\alpha}^{\beta} u_{\alpha}^{\beta} u_{\alpha}^{\gamma} + O(u^{4}) , $$

where $U_{0}$ is the equilibrium potential energy, $u_{\alpha}^{\beta}$, $u_{\alpha}^{\beta}$, and $u_{\alpha}^{\beta}$ are the displacements of $i$-th atom in $\alpha$ direction, $j$-th atom in $\beta$ direction, and $k$-th atom in $\gamma$ direction, respectively. $\Phi_{\alpha}^{\beta}$ are the second-order
IFCs and $\Psi_{ijkl}^\alpha$ are the third-order IFCs. $O(u')$ represents the higher-order terms. The force acting on each atom is $\mathbf{F}_i = -\nabla U$, and if all the atoms are at equilibrium state, $\mathbf{F}_i = 0$, so there is no first-order term in Eq. (13). Under harmonic approximation and only keeps the second-order term, we can easily get $F_i^\alpha = -\sum_j \Phi_{ijkl}^\alpha u_j$. If the $i$-th atom in the crystal is the $b$-th atom in the $l$-th unit cell and atomic index $j$ corresponds to $(i',b')$ in a similar convention, one can get the following equation from Newton’s second law of motion

$$m_i \ddot{u}_i^\alpha(t) = -\sum_{j\neq i} \Phi_{ijkl}^\alpha u_j^\alpha(t),$$

where $m_i$ is the mass of the $b$-th atom. $\ddot{u}_i^\alpha(t)$ is the acceleration of the atom $(i,b)$ in $\alpha$ direction at time $t$. Plane wave solutions for the equation above can be found as

$$u_i^\alpha(t) = \frac{1}{\sqrt{m_i \lambda_i}} e^{i \omega_i \lambda_i t} \lambda_i \epsilon_i e^{i \left( R_i - R_{i,b} \right)},$$

where $\epsilon_i$ is the eigenvector of $b$-th atom for phonon mode $\lambda_i$, $\lambda_i$ is the frequency of the wave, $R_i$ is the lattice vector for the $l$-th unit cell. It should be noted that in this equation $i$ is the imaginary unit, not the atomic index. By substituting Eq. (15) into (14) we can get the following equation

$$\omega_i^2 \epsilon_{i,b} = D_{i,\alpha}^\alpha \left( q \right) \epsilon_{i,b},$$

where $D$ is the “D-type” dynamical matrix that is expressed as

$$D_{i,\alpha}^\alpha \left( q \right) = \frac{1}{\sqrt{m_i m_j}} \sum_{l} \Phi_{ijkl}^\alpha \epsilon_{l,\alpha} e^{i \left( R_i - R_{i,b} \right)}.$$}

Solving Eq. (16) one can obtain the phonon dispersion relation $\omega(q)$.

**Phonon scattering mechanism and relaxation time**

As shown in Fig. 4, the scattering term in BTE includes phonon-phonon scattering, phonon-boundary scattering, and phonon-impurity (isotope, defects) scattering, etc. It can be seen from Eq. (6) that the phonon relaxation time $\tau_\alpha$ is defined through the scattering term and we need to quantitatively analyze the scattering mechanisms in order to calculate $\tau_\alpha$.

1) **Phonon-phonon scattering**

As shown in the lattice dynamics method, phonons are analyzed with the plane wave model. From wave-particle duality, we know that wave can also be described as particle. By analogy with the photons of the electromagnetic field, the quanta of the lattice vibrational field are referred to as phonons, which follow the quantum-mechanical rules. With quantum-mechanical methods, the Hamiltonian for the crystal with potential energy $U$ is in the form of

$$\hat{H} = \sum_i \frac{\hat{p}_i^2}{2m_i} + \frac{1}{2} \sum_{i,j} \Phi_{ijkl}^\alpha \hat{u}_i^\alpha \hat{u}_j^\alpha + \frac{1}{3} \sum_{i,j,l} \Phi_{ijkl}^{abc} \hat{u}_i^a \hat{u}_j^b \hat{u}_l^c + O(u'),$$

where the atomic index $i$ corresponds to $(i,b)$, $j$ corresponds to $(i',b')$, and $k$ corresponds to $(i''b'')$. $\hat{p}_i$ is the momentum operator for the $b$-th atom in $l$-th unit cell. Here, we only discuss the three-phonon scattering processes since it is the dominating process and prevails in literature. (Note that the theoretical framework of four-phonon scattering process has also been recently reported). The Hamiltonian can be separated into harmonic and anharmonic terms. In quantum mechanics, such a problem will be first solved with the harmonic oscillator model and then analyzed by adding the anharmonic perturbation.

We can introduce the intrinsic transition probability term $O_{\text{int}}$ for simplicity in the following derivation

$$Q_{\text{int}}^{ab} = \frac{\hbar \pi}{4N_0} \sum_{\mathbf{q} \neq 0} \Phi_{ijkl}^\alpha \Phi_{ijkl}^\beta \frac{e^{i \mathbf{q} \cdot \left( \mathbf{R}_i - \mathbf{R}_j \right)}}{\sqrt{m_i m_j m_b m_c}} \left[ \delta(n_e + n_b - n_a - n_c) \right],$$

where $\Lambda$ denote Kronecker delta function and $\delta$ designates the Dirac delta function. With Fermi’s golden rule, the scattering probability for the process satisfying $\omega_a + \omega_b = \omega_c$, can be derived as

$$\left( \frac{\partial \omega_c}{\partial t} \right)_\text{scat} = \sum_{\mathbf{q} \neq 0} \left[ -n_e n_b(1 + n_a) + (1 + n_b)(1 + n_e)n_a \right] Q_{\text{int}}^{ab}.$$  

Similarly, we can express the scattering probability for processes satisfying $\omega_a = \omega_c + \omega_b$, as

$$\left( \frac{\partial \omega_c}{\partial t} \right)_\text{scat} = \sum_{\mathbf{q} \neq 0} \left[ -n_b(1 + n_e)(1 + n_a) + (1 + n_e)(1 + n_b)n_a \right] Q_{\text{int}}^{ab}.$$  

The total scattering probability for three-phonon scattering processes is then

$$\left( \frac{\partial \omega_c}{\partial t} \right)_\text{scat} = -\sum_{\mathbf{q} \neq 0} \left[ n_e n_b(1 + n_a)(1 + n_b) - (1 + n_e)(1 + n_b)n_a \right] Q_{\text{int}}^{ab}.$$  

This equation will be used to calculate relaxation time due to phonon-phonon scattering.

2) **Other scattering mechanisms**

The other scattering mechanisms except phonon-phonon scattering include phonon-impurity scattering, phonon-boundary scattering, phonon-electron scattering, phonon-grain boundary scattering, and phonon-defect scattering, etc. There have been empirical equations to consider phonon-impurity scattering and phonon-boundary scattering by treating them as relaxation times. The scattering rate due to phonon-impurity scattering is shown as

$$\frac{1}{\tau_\beta} = \frac{\pi \alpha_\beta}{2N_0} \sum_{\mathbf{q} \neq 0} \delta(\omega_b - \omega_\beta) \sum_{\mathbf{b}} g(\mathbf{b}) |\mathbf{\epsilon}_{\mathbf{b},\beta}|^2,$$

where $g(\mathbf{b}) = \sum_{\mathbf{r}} f_r(\mathbf{b}) \left( \frac{1}{\sqrt{m_b}} \right)^2$. $f_r$ is the concentration of impurity atom s. The average mass is calculated with $m_b = \sum_{\mathbf{r}} f_r(\mathbf{b}) m_{b,r}$. $N_b$ is the number of wave vector $\mathbf{q}$ points.

The following equation is used to calculate the relaxation time due to boundary scattering

$$\tau_{\text{b}} = \frac{1 + p L_b}{1 - p |\mathbf{b}|} \tau_\alpha,$$  

where $p$ is the specularly parameter of the boundary and $L_0$ is the effective boundary mean free path in the Casimir limit. We need to emphasize here that phonon boundary scattering is a surface phenomenon, which is different from other scattering mechanisms, which are bulk behavior. If the phonon boundary scattering is regarded as a relaxation time, it actually assumes that the boundary scattering can be averaged as a bulk scattering process, which is a significant assumption. Rigorous treatment of phonon-boundary scattering requires solving the phonon BTE with proper boundary conditions and will be discussed later in Sec. 6.
SMRTA and iterative solution.

With non-equilibrium introduced by a small temperature gradient, the phonon distribution function will deviate from equilibrium distribution,

\[ n_i = n_i^0 + n_i^1. \]  

(25)

This applies to all the phonon modes. The deviation \( n_i^1 \) depends on the phonon distribution function of all other phonon modes. The single mode relaxation time approximation (SMRTA) method assumes that when we calculate the phonon distribution function for a specific mode, other phonon modes stay in their equilibrium states,

\[
\begin{cases}
  n_k = n_k^0 + n_k^1, \\
  n_i = n_i^0, \\
  n_{\gamma} = n_{\gamma}^0.
\end{cases} \]  

(26)

It means that the distribution of phonon mode \( \lambda \) deviates from equilibrium, while the other phonons are still in equilibrium. For processes with \( \omega_1 + \omega_2 = \omega_\gamma \), we can prove the following relation

\[(1 + n_i^0)(1 + n_k^0)n_{\gamma} = n_i^0n_k^0(1 + n_{\gamma}^0).\]  

Plug the Eq. (26.) into Eq. (20.) and use this relation to simplify the expression, we get

\[
\left( \frac{\partial n_i}{\partial t} \right)_{\text{accl}} = \sum_{\gamma} n_i^0 (n_k^0 - n_{\gamma}^0) Q_{\gamma,ik}^{(i)}.
\]  

(27)

For processes with \( \omega_1 + \omega_2 = \omega_\gamma \), we can show that the relation

\[(1 + n_i^0)(1 + n_k^0) = n_i^0(1 + n_k^0)(1 + n_{\gamma}^0) \]  

will be satisfied. We can then plug Eq. (26.) into Eq. (21.) and use this relation to simplify the expression to get

\[
\left( \frac{\partial n_i}{\partial t} \right)_{\text{accl}} = \sum_{\gamma} n_i^0 (1 + n_k^0 + n_{\gamma}^0) Q_{\gamma,ik}^{(i)}.
\]  

(28)

With Eq. (6.), (27.), and (28.), the relaxation time due to three-phonon scattering processes \( \tau_{3f} \) can be derived as

\[
\frac{1}{\tau_{3f}} = \sum_{\gamma} \left( n_{\gamma} - n_i^0 \right) Q_{\gamma,ik}^{(i)} + \frac{1}{2} (1 + n_k^0 + n_{\gamma}^0) Q_{\gamma,ik}^{(i)}.
\]  

(29)

If we further consider the phonon-impurity scattering and phonon-boundary scattering in the calculation, the total relaxation time can be calculated with Matthiessen’s rule\(^*\)

\[
\frac{1}{\tau_s} = \frac{1}{\tau_{3f}} + \frac{1}{\tau_{3f}} + \frac{1}{\tau_s}.
\]  

(30)

The superscript 0 in \( \tau_s^0 \) indicates the relaxation time calculated with SMRTA. This result can then be combined with Eq. (12.) to predict the lattice thermal conductivity.

With full iterative method, the deviation of distribution of all the three phonons are considered, as shown below

\[
\begin{cases}
  n_k = n_k^0 + n_k^1, \\
  n_i = n_i^0 + n_i^1, \\
  n_{\gamma} = n_{\gamma}^0 + n_{\gamma}^1.
\end{cases}
\]  

(31)

Since less assumptions are made, the iterative method can be more accurate than the SMRTA, especially for high thermal conductivity materials or at low temperature, when normal process of phonon-phonon scattering is important. Due to the page limit, the details of iterative solution will not be presented here. Interested readers are referred to related literature.\(^{34-36}\)

Implementation and application of first-principles BTE method

To implement the algorithm described above is quite challenging. The general workflow to calculate phonon thermal conductivity is show in Fig. 6. The first step is to extract the IFCs, from either classical potential or first-principles calculations. Since first-principles calculations can predict the ground state energy quite accurately without any input parameters, anharmonic lattice dynamics is usually combined with first-principles method. To accurately extract IFCs from first-principles simulation is not a trivial task. Currently, the most widely implemented approach to extract IFCs takes from atomic forces \( F(u) \) calculated from first-principles. One can generates atomic configurations by displacing certain atoms and using finite difference method to obtain IFCs.\(^\text{56} \)

Alternatively, one can randomly displace atoms to get atomic forces as a function of atomic displacements, and then use numerical fitting techniques, such as singular-value decomposition\(^\text{57} \) and compressive sensing\(^\text{58} \) to extract the IFCs. Both approach requires single-point DFT calculations of a large supercell for many configurations, which requires large computational cost. Another approach is to employ the linear response theory to directly obtain IFCs, also known as density-functional perturbation theory (DFPT).\(^\text{59} \) In all of these methods, considering the crystal symmetry is very important to reduce the computational cost. Another issue is the cutoff of IFCs. The IFCs theoretically should go over all atoms in the crystal. In reality, only the atoms that are close enough have a large enough IFC that should be considered. Therefore, a cutoff distance should be chosen and its effect on the lattice thermal conductivity should be carefully tested.\(^\text{60} \)

\[ \text{Fig. 6} \]  

The general numerical workflow for the first-principles BTE method. The harmonic and anharmonic IFCs are first extracted from the DFPT or the first-principles forces. The phonon specific heat and group velocity can be obtained from the harmonic lattice dynamics, and the scattering rates are extracted from the anharmonic IFCs. Then the thermal conductivity can be obtained by relaxation time approximation or the iterative solution of phonon BTE.

Due to the numerical uncertainty and the cutoff imposed to the
force constants, these raw IFCs from first-principles calculations generally do not satisfy invariance conditions\textsuperscript{59} that should be satisfied. The translational invariance condition has been proven to be significant to the accurate prediction of lattice thermal conductivity values, so it is usually artificially imposed by adding small compensation to each term in force constant.\textsuperscript{60} The thermal conductivity values can then be predicted using the anharmonic lattice dynamics method, with either SMRTA or the more accurate iterative method.

The first-principles method combined with anharmonic lattice dynamics is a very accurate method to predict the thermal conductivity without any fitting parameter. Many calculations based on this method has shown that the predicted thermal conductivity values agree well with experiments.\textsuperscript{77} Several sources of uncertainties can still be induced within the numerical calculation process. For example, the IFCs, especially the anharmonic IFCs can have some inaccuracy, due to the inaccuracy of first-principles calculation, choice of cutoff, and the finite difference method.\textsuperscript{11} The numerical integration within the first-Brillouin zone can also induce discretization error. Nevertheless, this is by far the most powerful tool that allows one to predict the lattice thermal conductivity of crystalline materials with the only information of initial atomic structures. On the other hand, it allows one to extract the thermal conductivity of each phonon mode (mode-resolve thermal conductivity). These results can be further combined with numerical solution of BTE to predict the lattice thermal conductivity of nanostructures.

To date, researchers from different groups have developed many different open source packages to implement the anharmonic lattice dynamics calculations, such as ShengBTE,\textsuperscript{36} AlmaBTE,\textsuperscript{44} phono3py,\textsuperscript{65} d3q,\textsuperscript{60} AFLOW-AAPL,\textsuperscript{66} etc. These packages are nicely interfaced with first-principles simulation tools, especially VASP and Quantum Espresso.

With this approach, Broido et al. first reported the full first-principles thermal conductivity prediction of Si and Ge without any adjustable parameters and the calculation results are comparable with experiments.\textsuperscript{25} Turney et al. extracted force constants from Lennard-Jones potential and predicted the thermal conductivity of Argon and compared the results with molecular dynamics simulations in 2009.\textsuperscript{25} This method has been applied to numerous materials, especially after 2015 when relevant software packages\textsuperscript{58} becomes publicly available. The readers are referred to some review articles, such as Ref. 46 and 47 on the application of the anharmonic lattice dynamics method to different materials. One of the most notable contribution is the successful prediction of the ultrahigh lattice thermal conductivity of cubic boron arsenide (BAs) in 2013, which is about 2240 W/mK at room temperature.\textsuperscript{57} The ultrahigh thermal conductivity is due to the large phonon band gap between acoustic and optical phonon branches, which significantly reduces the phonon scattering rates. Four-phonon scattering processes in BAs are further considered in 2017, and it reduces the room temperature thermal conductivity to about 1400 W/mK.\textsuperscript{59} Nevertheless, it is still among the most heat conductive semiconductor materials. Only very recently in 2018, high quality cubic BAs has been fabricated by three different groups and thermal conductivity is reported to be 1000–1300 W/mK.\textsuperscript{60} Since the results from the three different groups are similar, we select the results obtained by Kang et al\textsuperscript{66} and show in Fig. 7. The theoretical prediction and then experimental validation of a new material is a good example to demonstrate the significance of this powerful predictive tool.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image}
\caption{Fig. 7 The measured thermal conductivity of BAs and the predictions from first-principles PBTE approach considering different scattering mechanisms. One can see that if only three phonon scattering is considered, the predicted lattice thermal conductivity is much larger than the experimental value. If four phonon process is further included, the predicted value agree quite well with experimental value in a relatively broad temperature range. \copyright{} from Ref. 68 with permission from AAAS.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image}
\caption{Fig. 8 (a) The atomic structure of graphene and silicone. They have the same hexagonal lattice, but silicene has a small bulking and thus does not have reflectional symmetry in the z direction. (b) The predicted lattice thermal conductivity of graphene, adopted from Ref. 74. It can be seen that ZA modes have dominate contribution. Also, the SMRTA (denoted as RTA in the figure) predicted result is much smaller than the full iterative solution (denoted as Total). (c) The lattice thermal conductivity of silicene decomposed into different directions. It can be seen that the z displacement has a very small contribution to the lattice thermal conductivity, which is quite different from graphene. Reproduced from Ref. 74 and 83, with permission from American Physical Society.}
\end{figure}
Such a method has also been applied to low dimensional structures, including nanowire, thin films, graphene, and other 2D materials, etc. Here we just briefly review the thermal transport investigation of 2D materials. In 2010, Lindsay employed the optimized Tersoff potential to investigate the lattice thermal conductivity of graphene. It has been shown that iterative approach is very important to correctly predict the thermal conductivity of graphene, since the Normal process is important. The major contribution of out-of-plane vibrational modes (the flexural branch, or ZA mode) to the lattice thermal conductivity in graphene has also been identified. Due to the refleitional symmetry, the three-phonon scattering events involve odd number of ZA modes are forbidden, which give rise to a high ZA phonon lifetime. There are many subsequent calculations on graphene, employing IFCs from first-principles simulation, which basically confirms these findings. The predicted thermal conductivity of graphene is generally on the order of 3000 W/mK, which agree well with experimental results. Note that all these calculations only consider three-phonon process. A recent work using the empirical optimized Tersoff potential indicates that four-phonon scattering can reduce the contribution of out-of-plane mode to the thermal conductivity at room temperature.

As shown in Fig. 8, the thermal conductivity of silicene, a cousin of graphene, has also been investigated and the thermal conductivity is much lower, due to the bulking structure that breaks the reflectional symmetry. The thermal conductivity of other 2D materials have also been investigated, including metal dichalcogenide and black phosphorene etc. Unlike bulk materials, less consistency has been found in the reported thermal conductivity values of 2D materials. There has been a few possible reasons. First, although there were debates, now it becomes clear that the flexural branch should have a quadratic dispersion at zero strain and becomes linear when an infinitesimal strain is applied. However, it is very difficult to obtain the quadratic dispersion for two reasons. First, it is difficult to avoid strain in first-principles simulations. Second, the numerical uncertainty of harmonic FCs makes it difficult to satisfy invariance conditions, and the rotational invariance is shown to be important to obtain a quadratic dispersion. The inaccuracy of anharmonic IFCs could be another issue for the inconsistency. The choice of cutoff distance for IFCs should be carefully tested when considering the anharmonic IFCs and the acoustic sum rule should be properly applied. It has also been shown that the intrinsic force inaccuracy from first-principles calculation could also affect the correct prediction of thermal conductivity of 2D materials. In short, the first-principles BTE method should be carried out with care when applied to 2D materials.

Another important aspect of this method is that it can obtain the thermal conductivity of each phonon mode, or the mode-resolved thermal conductivity. This allows us to construct the wavelength accumulation function and the mean free path accumulation function of thermal conductivity, as shown previously in Fig. 3. This information guides the material design and allows us to better engineer the lattice thermal conductivity of solid materials. It can also serve as an input to the numerical solution of BTE, as will be discussed later in Sec. 6.

### 3.3 Electron-phonon coupling in metal and doped semiconductors

As shown in Fig. 2, in metals and doped semiconductors, heat is carried by both electrons and phonons, and hence the total thermal conductivity can be decomposed into the electron and phonon parts, $\kappa = \kappa_e + \kappa_p$. Electronic contribution dominates the heat transfer in metals, but phonon contribution can also be important. In metals, temperature gradient drives the transport of both electrons and phonons. Electron-phonon scattering is the limiting factor of the transport process. Electron scattering by phonons is the major reason for finite MFP of electrons. On the other hand, it also adds one additional scattering channel for phonons. The electron-phonon scattering can be understood from different pictures: one can regard electrons and phonons as particles and they scatter with each other during transport process; alternatively, electron-phonon coupling can be regarded as the perturbation of electronic eigenstate (wave function) under the lattice vibration.

The phonon-electron ($p$-$e$) scattering rate can be obtained from Fermi’s golden rule (FGR). The electron-phonon ($e$-$p$) matrix element, which describes the rate of an electron at initial state, $\langle \mathbf{k} | \mathbf{v} \rangle$ is scattered to $i | \mathbf{k} + \mathbf{q} \rangle$ by a phonon mode $\lambda = (\mathbf{q}, \nu)$, is defined as

$$g^{\lambda}_{i \rightarrow q} = \frac{\hbar}{2\omega_{\lambda}} \langle \mathbf{v}_{\lambda} \rangle |\mathcal{U}^f_{\lambda, i} |^2 ,$$

(32)

where $\psi$ is the ground-state Bloch wave function, $\mathcal{U}^f_{\lambda, i}$ is the first-order derivative of the Kohn-Sham potential with respect to the phonon displacement, $\mathbf{k}$ is the wave vector of the initial electron state, and $i, j$ denote the electronic band indices of the initial and final states, respectively. Note that due to the momentum conservation requirement, the final state must have a wave vector of $\mathbf{k} + \mathbf{q}$. This parameter can be extracted in the DFPT calculations from Wannier function interpolation. The accurate calculation of $e$-$p$ scattering matrix element requires deep understanding of the density functional theory and solid state physics, which used to be a challenging task. Fortunately, such calculations were implemented in first-principles packages very recently, such as Quantum Espresso with EPW (Electron-Phonon coupling using Wannier functions). As such, if one is only interested in the transport properties of solids, one can start with the $e$-$p$ matrix elements, which can be directly obtained from the existing open source code.

Under the relaxation time approximation, the scattering rate (i.e. the inverse of relaxation time $\tau$) of phonon mode $\lambda$ is the sum of all the possible phonon-electron scattering processes, $i.e.$

$$\frac{1}{\tau_{\lambda}} = \frac{2\pi}{\hbar} \sum_{\mathbf{q}, \nu} g^{\lambda}_{i \rightarrow q} \delta \left( f_{\lambda} - f_{\lambda + \mathbf{q} + \hbar \omega_{\lambda}} \right) ,$$

(33)

where $g$ is the $e$-$p$ interaction matrix element, $f$ is the Fermi-Dirac distribution function, $i$ and $j$ are electron band indices, $\nu$ is the electron energy, and $\omega_{\lambda}$ is the phonon frequency. $\delta$ function here is to ensure the conservation of energy during the scattering process. Note that to accurately calculate the scattering rate, Brillouin zone integration with very fine $\mathbf{k}$ mesh and $\mathbf{q}$ mesh is required. With the obtained phonon scattering rate by considering phonon-electron ($p$-$e$) and phonon-phonon ($p$-$p$) scattering process, the total phonon relaxation time can be obtained using Matthiessen’s rule as following

$$\frac{1}{\tau_{\lambda}} = \frac{1}{\tau_{\lambda}^{ee}} + \frac{1}{\tau_{\lambda}^{pp}} .$$

(34)

With the phonon-phonon scattering rate $\tau_{\lambda}^{pp}$ calculated from the anharmonic lattice dynamics calculations as described above in Sec. 3.2, the total scattering rate of phonons can be calculated. Then combining with Eq. (12.), one can calculate the lattice thermal conductivity of metals or doped semiconductors.
To further obtain the electron contribution, the Onsager relations, derived with electron BTE, can give the electronic transport properties\(^8\):

\[
\sigma_{e-p} = -\frac{e^2 n}{\epsilon} \sum \frac{\partial f_k}{\partial \epsilon} \epsilon_{n_k} v_{n_k} \tau_{n_k} ,
\]  
\[
(35)
\]

\[
\begin{align*}
[\sigma_S] = & -\frac{e m}{\epsilon T} \sum \left( \epsilon_{n_k} - \mu \right) \frac{\partial f_k}{\partial \epsilon} \epsilon_{n_k} v_{n_k} \tau_{n_k} ,
\end{align*}
\]

\[
(36)
\]

\[
K_{e-p} = -\frac{n e}{\epsilon T} \sum \left( \epsilon_{n_k} - \mu \right) \frac{\partial f_k}{\partial \epsilon} \epsilon_{n_k} v_{n_k} \tau_{n_k} ,
\]

\[
(37)
\]

where \( \Sigma, S \) and \( K \) are \( 3 \times 3 \) tensors. \( \Sigma \) is the electrical conductivity, \( S \) is the Seebeck coefficient, and \( K \) is related to the electron thermal conductivity \( k_e = K \cdot S_0 \cdot S \), where \( T \) is the temperature. The summation in these three equations is over all the electrons enumerated using electronic wave vector \( k \) and band index \( i \). The \( e \) is the elementary charge, \( n \) is the number of electrons per state, \( V \) is the volume of primitive cell, \( f_k \) is the Fermi-Dirac distribution, \( \epsilon_k \) is the electron energy, \( \mu \) is the chemical potential, \( v_k = \frac{1}{h} \frac{\partial \epsilon_k}{\partial k} \) is the electron velocity, \( \alpha \) and \( \beta \) denotes the directional component, and \( \tau_k \) is the electron transport relaxation time. The electron transport relaxation time, limited by \( e-p \) scattering, can be obtained by considering the electron-phonon interactions as\(^9\)

\[
\frac{1}{\tau_k} = 2\pi \hbar \sum_{\lambda} \sum_{\epsilon_{n_k}} \delta_n \left( \frac{1}{F_{\lambda}} \right) \delta \left( \epsilon_{n_k} + \hbar \omega - \epsilon_{\lambda} \right) \frac{\partial f_k}{\partial \epsilon} \left( \epsilon_{n_k} - \hbar \omega, \epsilon_{\lambda} \right) \frac{1}{\tau_{\lambda}} \left( \frac{1 - \frac{v_{\lambda} \tau_{\lambda}}{1 - \frac{v_{\lambda}}{\tau_{\lambda}}} \right) .
\]

\[
(38)
\]

**Implementation and applications**

To calculate the aforementioned quantities and obtain the electron and phonon thermal conductivity of metals, one first need to perform first-principles calculation to obtain the electron band structure \( \sigma(k) \) and wave functions, which are direct output of DFT calculations. For phonon thermal conductivity, the phonon-phonon and phonon-electron scattering are considered. The phonon dispersion curve and phonon relaxation times (due to phonon-phonon scattering only) should be calculated using harmonic lattice dynamics, which is the same as non-metals. Moreover, the phonon property should be combined with Kohn-Sham wave functions to further calculate the electron-phonon scattering matrix elements. It should be noted that the electron-phonon scattering matrix elements are initially obtained on coarse electron and phonon wavevector grids and then interpolated to finer electron and phonon wavevector grids using the maximally localized Wannier functions basis as implemented in the electron-phonon Wannier (EPW) package.\(^{90,91}\)

Then, proper Brillouin zone integration needs to be performed to obtain the phonon relaxation time due to phonon-electron scattering, and the electron relaxation time due to electron-phonon scattering. Note that the Brillouin zone integration is not trivial, proper choice of discretization mesh is important to obtain correct and accurate results. Now most of the calculations can be carried out with the EPW package.\(^{94,95}\)

The full first-principles calculation of thermal property of noble metals are reported in 2016.\(^96\) The lattice thermal conductivity values at room temperature are found to be 6, 4, and 2 W/mK respectively for Al, Ag, and Au. The electron thermal conductivities are 246, 370, and 276 W/mK, respectively. The predicted results of total thermal conductivity agree well with experimental values.\(^97\) Wang et al.\(^7\) also calculated the lattice thermal conductivity of more metals, including Cu, Ag, Au, Pt, and Ni. It was found that the effect of electron-phonon coupling to lattice thermal conductivity is found to be important in Pt and Ni than in other materials. In 2017, Tong and Bao\(^5\) calculated the electron and phonon thermal conductivity for intermetallic compound NiAl and Ni3Al, the reported value also agree very well with experiment, as shown in Fig. 9. It is found that the electron-phonon coupling is also important in both materials. Lattice thermal conductivity is more important at lower temperature regime, and its effect decays as temperature is higher.

In addition to metals, the electron-phonon coupling calculation is also important in other aspects of thermal transport. For example, in heavily doped semiconductors, it has been reported that the

![Fig. 9 Variation of phonon thermal conductivity (\( k_p \)), electron thermal conductivity (\( k_e \)) and total thermal conductivity (\( k_{\text{total}} = k_p + k_e \)) with temperature for (a) NiAl and (b) Ni3Al. Reproduced from Ref. 98, with permission from Elsevier.](image-url)
Phonon-electron scattering can significantly reduce the lattice thermal conductivity.\cite{1,26-30} Note that in doped semiconductors, the doping atoms themselves also affect thermal conductivity through phonon-impurity scattering, but the results only considered the phonon scattering with the charge carriers induced by the dopants. Therefore, the calculated results cannot be directly compared with experimentally measured thermal conductivity data for heavily doped semiconductors. Another important application of this method is to model the non-equilibrium process in energy carrier dynamics. For example, Vallabhaneni et al.\cite{37} calculated electron-phonon scattering and combined with multi-temperature model, to simulate the thermal measurement of graphene using Raman spectroscopy. It was found that non-equilibrium in phonon distribution will affect the temperature measurement. The Raman spectra reflect the optical phonon temperature, while the more thermally conductive acoustic phonons have lower temperature. The prediction has been validated by experiment recently.\cite{99}

### 4. Molecular dynamics simulation

The molecular-dynamics (MD) simulations technique can provide a classical description of the dynamic evolution process of the N-atom system.\cite{102,103,104} During the MD simulations, the information of each atom, including atomic position $x$, velocity $v$, and force $F_x$, is calculated at each step, which is then used to predict the properties in the next step. In the dynamic process, the motion of each particle is described by the Newton's second law:

$$m_i \frac{d^2x_i}{dt^2} = \sum_{j=1 \neq i}^{N} F_{x_j} = F_i,$$

(39)

where $m_i$ is the mass of atom $i$, and $F_{x_j}$ is the force exerted by atom $j$ on atom $i$. The force term is derived from the interatomic potential that must accurately fit the potential energy surface. Obviously, the interatomic potential greatly affects the accuracy of MD simulations. The commonly used empirical potentials are pairwise interactions, such as Lennard-Jones potential and Morse potential. In addition, other potentials also include three-body interaction for structure stability, such as Stillinger-Weber potential and Tersoff potential. It should be noted that as a classical system, all the modes (degrees of freedom) are equally excited in MD simulation. However, at low temperature, high-frequency modes are frozen in the quantum system, which deviates from the classical case. Quantum corrections\cite{31} are used to qualitatively account for this discrepancy.

Phonons are the main heat carrier in semiconductors and insulators, and the dynamics of lattice vibration can be well captured in MD simulation at evaluated temperature. Furthermore, the MD simulation has the advantages of implicitly modeling the anharmonic phonon-phonon interaction to all orders, compared to the limitation of handling only three-phonon or four-phonon scatterings in most first-principles BTE calculations.\cite{32,33} Moreover, MD simulations are also effective to study various realistic effects on thermal transport, including folding,\cite{34} strain,\cite{35} defects,\cite{36,37} dislocation,\cite{38} grain boundary,\cite{39} doping,\cite{40} and so on. In addition to the prediction of thermal conductivity, MD simulations are also widely used to analyze the thermal transport behaviors and provide valuable insight to the understanding of underlying mechanisms.

#### 4.1 EMD and NEMD simulations

There are two commonly used methods to calculate the thermal transport properties, for instance lattice thermal conductivity, based on MD simulations.\cite{41,42} The first method uses equilibrium MD (EMD) simulation, which calculates the thermal conductivity based on the fluctuation-dissipation theorem using Green-Kubo formula. Thus, this method is also known as Green-Kubo method. The second method is based on the Fourier's law of heat conduction, in which the non-equilibrium temperature distribution and heat flow are directly monitored under a temperature gradient. Therefore, this method is conducted under a temperature bias in the non-equilibrium MD (NEMD) simulation, and is also known as the direction method. In the next section, these two methods are described in details.

**EMD simulations**

**Thermal conductivity** The Green-Kubo method computes lattice thermal conductivity based on the fluctuation-dissipation theorem\cite{43} of heat current in EMD simulations. In details, the lattice thermal conductivity can be computed from the heat current auto-correlation function (HCACF) based on Green-Kubo formula as

$$\kappa_{\alpha\beta} = \frac{1}{k_B T^2 V} \langle \frac{\partial}{\partial t} \langle J_\alpha(0) J_\beta(t) \rangle \rangle dt,$$

(40)

in which $J_{\alpha}$ is the heat current in the $\alpha$th direction, $k_B$ is the Boltzmann constant, $T$ is system temperature, and $V$ is the system volume. The angular bracket means the ensemble average. One advantage of Green-Kubo method is that one can get the thermal conductivity tensor along different directions with the same set of EMD simulation results. For isotropic bulk materials, thermal conductivity is usually averaged over three diagonal components so that a factor of three will appear in the denominator of Eq. (40). For one-dimensional materials, such as the carbon nanotube, HCACF is computed along the axial direction.

In EMD simulation, the heat current is defined as\cite{44}

$$J(t) = \frac{d}{dt} \sum_i x_i(t) F_{x_i}(t),$$

(41)
where $E_i(t)$ and $\mathbf{x}_i(t)$ are the time-dependent total energy and coordinate of the $i$th atom, respectively. The expression of heat current can have complex form, depending on the specific terms in the interatomic potential (e.g., many-body potential). Take the widely used pairwise potential for instance, the heat current in Eq. (41) can be written as:

$$J = \sum_i E_i \mathbf{v}_i + \frac{1}{2} \sum_i \sum_j \left[ F_{ij} \cdot (\mathbf{v}_i + \mathbf{v}_j) \right] (\mathbf{x}_i - \mathbf{x}_j), \quad (42)$$

where $F_{ij}$ is the interatomic force acting on atom $i$ due to the pairwise interaction with atom $j$. In Eq. (42.), the first term describes the convection contribution, and the second term is the conduction part. Benefit from pairwise empirical potential in MD simulation, the pair force $F_{ij}$ in Eq. (42.) can be easily calculated from the derivative of the pair potential. Moreover, when considering many-body interactions, such as Stillinger-Weber potential, extra terms containing three-body force $F_{ijk}$ should also be included in the expression of heat current.\(^{115}\) Different from the kinetic theory that is only applicable to crystalline solids, the Green-Kubo method has been widely used to calculate thermal conductivity of various materials, including disordered systems in which the phonon gas picture is invalid.\(^{120}\)

The main challenge in EMD based thermal conductivity calculations arises from the difficulty to carry out the time integral up to infinity and ensure its convergence. There are mainly two types of implementations of the Green-Kubo method in literature: one is the time-domain approach, and the other is the frequency-domain approach.

The time-domain approach is to handle HCACF in time-domain. The most straightforward way is the direct integration method, which replaces the integral with summation and numerically records HCACF in time-domain as:

$$\text{Cor}(t) = \left\langle J_\alpha(0) J_\beta(t) \right\rangle = \frac{1}{N - m} \sum_{n=0}^{N-m} J_\alpha(n t_0) \left\langle (n + m) t_0 \right\rangle, \quad (43)$$

where $t_0$ is the time step, $N$ is the total number of time steps recorded in EMD simulations, and $m = 1/t_0$ is the integer number for time $t$. Furthermore, the infinite time integral is replaced by a summation up to a finite cut-off time $t_c$,

$$K_{\alpha\beta} = \frac{t_0^2}{k_B T} \sum_{t=t_0}^{t_c} \text{Cor}(t). \quad (44)$$

In this direct integration method, the finite cut-off time $t_c$ in the integration should be long enough to ensure the proper decay of HCACF. For a given finite time EMD simulation, first avalanche method has been proposed to accurately determine the cut-off time based on the statistical analysis of HCACF.\(^{117}\)

Another widely used approach is based on the relaxation time fitting.\(^{117,119-122}\) Because of the different relaxation times for acoustic and optical phonons, HCACF generally exhibits a two-stage decaying behavior, with a very fast acoustic decay at the beginning and followed by a much slower optical decay. Thus, HCACF can be fitted by the double exponential function.\(^{122}\) In consideration of the large relaxation time of long wavelength phonons and the finite cut-off time (as shown in Fig. 11), Chen et al.\(^{117}\) proposed a nonzero correction to this fitting method,

$$\frac{\text{Cor}(t)}{\text{Cor}(0)} = A_1 e^{-\tau_1 t} + A_2 e^{-\tau_2 t} + Y_0, \quad (45)$$

where $A_1$ and $A_2$ are two fitting parameters, $\tau_1$ and $\tau_2$ are the relaxation time obtained from the fitting, and $Y_0$ is a tiny nonzero value for correction. Correspondingly, the thermal conductivity can be calculated as

$$K_{\alpha\beta} = \frac{\text{Cor}(0)}{k_B T} \left( A_1 \tau_1 + A_2 \tau_2 + Y_0 \tau_0 \right). \quad (46)$$

The nonzero correction term accounts for the contribution from long wavelength phonons that have relaxation time longer than the cut-off time.

The frequency-domain approach is to handle the time integral indirectly via Fourier transform. For instance, the bulk thermal conductivity from Green-Kubo formula can be viewed as the Fourier transform of HCACF, and one can define the spectral thermal conductivity as

$$K(\omega) = \frac{1}{3k_B T^3} \int_0^{t_c} \langle J(0) \cdot J(t) \rangle e^{i \omega t} dt. \quad (47)$$

Fig. 11 The calculations of thermal conductivities from EMD and NEMD simulations. (a) The normalized heat current autocorrelation function from EMD simulation for raw data (red line) and double exponential fitting (blue line) before the cut-off time in a $4 \times 4 \times 4$ supercell for bulk silicon. The inset of (a) show the long time region near the cut-off time. (b) Temperature distribution in a hexagonal nitride boron ribbon from NEMD simulation. The inset of (b) shows the thermal conductivity of infinite system by using the extrapolation method. Reproduced from Ref. 117, with permission from Elsevier.
In this way, the conventional thermal conductivity is simply the spectral thermal conductivity in the static limit \((\omega \to 0)\). Based on the Wiener-Khinchin theorem, the spectral heat current power density can be expressed as

\[
|J(\omega)|^2 = \int_0^\infty \langle J(0) \cdot J(t) \rangle e^{i\omega t} dt,
\]

where \(J(\omega)\) is the spectral heat current defined as \(J(\omega) = \int_0^\infty J(t) e^{i\omega t} dt\). As a result, thermal conductivity can be calculated as

\[
K = K(\omega) \bigg|_{\omega=0} = \frac{|J(\omega)|^2}{3k_B T^2}.
\]

Based on the relaxation time approximation, Volz et al. proposed to fit the spectral thermal conductivity as

\[
K(\omega) = \frac{K(0)}{1 + i\omega \tau_0},
\]

in which \(K(0)\) and \(\tau_0\) are the fitting parameters, corresponding to the static thermal conductivity and single exponential decay constant.

Moreover, the bulk thermal conductivity from EMD simulations also exhibits the finite size dependence due to the use of periodic boundary condition. Therefore, in most of practical calculations, in addition to the autocorrelation time, the convergence on simulation domain size should be carefully tested and verified.

The EMD simulation method has been widely used to predict the lattice thermal conductivities of various materials. For example, Volz et al. used EMD method to predict the thermal conductivity of bulk silicon crystals and silicon nanowires. The suppressed thermal conductivity in silicon nanowires and temperature dependence is also observed from the EMD calculations. In addition, further engineering of thermal conductivity in silicon crystals are widely studied by using EMD method, such as silicon nanotube, and Si-Ge core-shell nanowires. The results show that the structural engineering can induce significant phonon scatterings, which further reduces the thermal conductivity dramatically.

On the other hand, the EMD method is also widely used to study the thermal conductivity of novel bulk materials that with accurate interatomic potential, such as carbon-based materials, PbTe, new two-dimensional materials, etc. For example, by using EMD method, Ye et al. reported that thermal conductivity of body-centered tetragonal \(C_4\) reaches ~1200 \(W m^{-1}K^{-1}\) in the cross-plane direction at room temperature, and is highly anisotropic. Carbogno et al. computed thermal conductivity of carbon schwarzites and found that the Gaussian curvature has a significant impact on thermal transport.

Interfacial thermal resistance In addition to the thermal conductivity, the classical EMD simulation has also been applied to calculate the interfacial thermal resistance between different materials based on temperature fluctuations. Rajabpour and Volz demonstrated that the interfacial thermal resistance can be defined as

\[
R = \frac{A}{k_B} \int_0^\infty \frac{(\Delta T(0) \Delta T(t))}{\left\langle \Delta T(0)^2 \right\rangle} dt \left( \frac{1}{N_1} + \frac{1}{N_2} \right),
\]

where the subscripts 1, 2 represent the two interacting surfaces around interface, \(\Delta T\) is the temperature difference between the interacting atoms of the two surfaces, \(A\) is the interfacial area, and \(N_1, N_2\) is the number of degrees of freedom at each side. During the EMD simulation, the temperature fluctuation at the interface is recorded as a function of time. Then, the interfacial thermal resistance is calculated based on Eq. (51). By using this EMD method, Rajabpour and Volz further demonstrated the existence of a universal interfacial thermal resistance at high frequency, which only depends on the frequency and a set of degrees of freedom of the two interacting bodies. Ni et al. has used this EMD method to study the inter-layer resistance in few-layer graphene.

First-principles molecular dynamics: The accuracy of classical MD simulations depends critically on the empirical potential used in the simulation. However, it is a challenging task to accurately fit the interaction potential surface for complex structures with empirical potential, and in many cases no adequate potential exists in literature. Therefore, it is high desirable to combine with MD simulations the first-principles \((ab\ initio)\) methods, for instance the density functional theory (DFT), which can provide interatomic interaction with high-accuracy.

Recently, some efforts are devoted to study the thermal conductivity from the \(ab\ initio\) EMD simulations. Marcolongo et al. proposed a technique to calculate the thermal current from \(ab\ initio\) simulations. However, their approach requires the solution of a set of linear equations at every MD step, which could be computationally expensive. Carbogno et al. introduced a quantum mechanical definition of a local stress tensor for the Cartesian components \(ab\) of a given ion as

\[
s_n = \frac{Z}{V} \left( \int d\mathbf{x} n(x)(\sum_{i} x_i^2 - x_i^2 \hat{\mathbf{x}} \cdot \hat{\mathbf{x}} - \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}^2) - \frac{1}{k_B T} \sum_{j} Z_j (\sum_{i} x_i^2 - x_i^2 \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}^2 - \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}^2) \right).
\]

in which the first term is the contribution from the interaction with their ground-state density \(n(x)\), and the second term is the electrostatic repulsion between the nuclei with charges \(Z\). In their formulism, the heat current can be calculated as

\[
\mathbf{J} = \sum_{i} E_i \mathbf{v}_i + \frac{1}{2} \sum_{i} \sum_{j \neq i} (F_{ij} \cdot \mathbf{v}_i)(x_i - x_j) = \sum_{i} E_i \mathbf{v}_i + \sum_{i} s_j \cdot \mathbf{v}_i.
\]

Carbogno et al. applied this method to calculate the thermal conductivity of bulk silicon and \(ZrO_2\), and obtained good agreement with the experimental results. However, the convective term of the heat current was ignored in their calculations, which can be important for liquids. Besides, the stress tensor expression is most suitable for an all-electron Hamiltonian, but not for other widely used \(ab\ initio\) calculations. There are also other efforts to develop first-principles Green-Kubo method for calculating thermal conductivity of liquids and amorphous solids.

NEMD simulations

Thermal conductivity In non-metallic solids, the transport of phonons, as the dominant heat carriers, is usually diffusive and can be described by the Fourier's law of heat conduction as

\[
K = -\frac{J}{\nabla T}.
\]

where \(J\) and \(\nabla T\) are the heat flux and temperature gradient along the transport direction in the non-equilibrium steady state. The local temperature can be computed from MD simulation as
\[ T = \left( \sum_{i=1}^{N} m_i \mathbf{v}_i \cdot \mathbf{v}_i \right) / (3Nk_B) , \] (55)

where \( N \) is the number of atoms in the local bin, and the angular bracket denotes the ensemble average. One key step in NEMD simulations is to build up the temperature gradient along the transport direction. There are two different ways to establish the temperature gradient: (i) impose the local temperature by thermostats (heat source/sink); (ii) impose the heat flux by simultaneously extracting and adding kinetic energy at heat bath.

Thermostat, also known as heat bath, is used in MD simulation to maintain the constant temperature, i.e., the canonical ensemble. There are two representative thermostats widely used in MD simulation. One is the Nosé–Hoover thermostat, in which a deterministic damping term is introduced into the equation of motion as

\[ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} - \zeta \mathbf{p}_i , \] (56)

Here, \( H \) is the Hamiltonian of the system, \( p_i \) is the momentum of the \( i \)th atom, and \( \zeta \) is a damping term determined by the dynamics of the heat bath atoms as

\[ \frac{d\zeta}{dt} = \frac{1}{T^2} \left( \sum_i N_i p_i \cdot p_i - 1 \right) , \] (57)

where \( T \) is the aimed temperature, \( N \) is the number of atoms in the heat bath, and \( T \) is the response time of the heat bath. The other is the Langevin thermostat, in which a random force \( \xi \) and the dissipation rate \( \lambda \) are introduced simultaneously as

\[ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} + \xi - \lambda p_i , \] (58)

\( \xi \) and \( \lambda \) is related to each other according to the fluctuation-dissipation theorem, which gives the variance of the random force as \( 2m\kappa T \).

The choice of thermostats may have significant impacts on the MD simulations of heat conduction at nanoscale. Previously, Chen et al.\cite{14} demonstrated that, compared to Nosé–Hoover heat bath, Langevin heat bath is more efficient to generate a linear temperature profile with small boundary temperature jump due to its stochastic nature. Moreover, Langevin heat bath can produce reliable physical results in thermal rectification transport, which are consistent with experimental results in a large range of heat bath parameter.

Another way to establish the temperature gradient is to impose a heat flux across the system. This can be achieved by rescaling atomic velocities in the heat bath, or artificially interchanging the velocity of the hottest atom in the cold region with the velocity of the coldest atom in the hot region. The imposition of heat flux reverses the conventional cause (temperature difference) and effect (heat flow) picture of heat conduction, and thus is also known as the reverse NEMD approach. Compared to the thermostat method with targeted temperature control, the frequency and amount of energy exchange should be carefully verified in order to establish the reasonable temperature distribution.

By rescaling atomic velocities in the heat bath region at each MD time step, heat energy \( \Delta E \) is added to the heat source on one side, and the same amount of energy is removed from the heat sink on the opposite end. Generally speaking, as the atomic velocity is scaled by the same factor in the source/sink region, the net kinetic energy is increased/decreased by the same amount. Therefore, the heat flux can be calculated as

\[ J = \frac{\Delta E}{S \Delta t} , \] (59)

where \( S \) is the cross sectional area, and \( \Delta t \) is the time interval performing the velocity rescaling.

After reaching the non-equilibrium steady state, the temperature gradient can be established by fitting the linear region of temperature distribution, as shown in Fig. 11. In regard to the modeling setup for temperature distribution, there are also two kinds of boundary conditions: non-periodic and periodic. One can set the heat baths at the two ends of system with boundary atoms fixed to avoid atom ablation, which is the non-periodic condition. On the other hand, to avoid the fixed boundary effect on thermal transport, another setup with periodic boundary condition is also widely used to establish the temperature gradient, in which a projection of temperature distribution is obtained in the other half side.\cite{18} In this setup, a factor of two will appear on the denominator of Eq. (59.) for heat flux to account for the bi-directional heat flow. One has to use this setup in certain case where the periodic boundary must be used, for instance when treating the long-range electrostatic interaction. However, the use of periodic boundary condition comes at additional price that one has to double the size of simulation domain compared to the case with fixed boundary condition.

As the calculated sample length smaller than the phonon mean free path, the phonon-boundary scatterings would occur at the interfaces with the heat source/sink. As a result, the thermal conductivity will be limited by the simulation size, which is known as the Casimir limit. For bulk materials, the sample size is much greater than the phonon mean free path so that the size effect will vanish. In the finite-size simulations, the effective mean free path \( l_{eff} \) is determined by

\[ \frac{1}{l_{eff}} = \frac{1}{l_0} + \frac{1}{L} , \] (60)

where \( L \) is the length of the simulation domain, and \( l_0 \) is the mean free path for infinitely large system. Eq. (60.) suggests a linear relationship between \( 1/l_0 \) and \( 1/L \), so that the converged thermal conductivity for bulk system can be obtained by extrapolating the simulation domain size to \( 1/L=0 \), as shown in the inset of Figure 11.

In low-dimensional materials, however, recent studies\cite{19,20} found thermal transport does not obey the Fourier’s law due to the anomalous heat diffusion, leading to the peculiar size-dependent thermal conductivity. For instance, in one-dimensional materials, Maruyama\cite{21,22} firstly demonstrated by carrying out NEMD simulations that the thermal conductivity for small diameter (5.5) carbon nanotube does not converge to a finite value when increasing in tube length up to 404 nm, and a power law dependence \( \kappa \sim L^\alpha \) is observed. Then, the NEMD method is considered as an effective approach to study the length dependent thermal transport behavior in low-dimensional systems, which have attracted lots of research interests.\cite{23,24,25} Similar length dependence of thermal conductivity is also reported in silicon nanowires\cite{26,27} and single extended polymer chains\cite{28} by using NEMD method. For two-dimensional materials, NEMD simulations\cite{29} found a logarithmically divergent length dependence of thermal conductivity in single-layer graphene. Chen et al.\cite{30} also reported such length dependent thermal conductivity from NEMD simulations. Moreover, they further studies demonstrated that such length dependent thermal conductivity can be
further suppressed by the substrate coupling or defect scatterings. Compared to the amorphous SiO$_2$ substrate, the hexagonal boron nitride is found to be a more promising substrate for single-layer with high heat dissipation ability for the less suppressed intrinsic high thermal conductivity and length dependence.

There are also some works studied the thermal transport properties based on first-principles NEMD simulations. The basic idea follows the theory as discussed above. However, the constrained calculation ability of first-principles makes it very challenging to study size dependent thermal conductivity for complex materials at large scale.

**Interfacial thermal resistance** The NEMD simulation has also been widely used to study the interfacial thermal resistance between different materials, known as the Kapitza resistance. By establishing a temperature jump $\Delta T$ at the interface, the interfacial thermal resistance $R$ is calculated as

$$ R = \frac{\Delta T}{J}, \tag{61} $$

where $J$ is the heat flux through the interface. This method is widely used to investigate the thermal transport across different interfaces, including solid/solid, solid/liquid, solid/gas, and so on. However, when one side of the interface is metal, electrons can contribute to heat transport, and one has to consider other heat transport mechanism, such as electron-phonon interaction. In this situation, other theory should be involved.

For instance, by performing NEMD simulations, Alexeev et al. found that the Kapitza resistance between few-layer graphene and water is strongly dependent on the layering of water adjacent to the graphene layers, exhibiting an inverse proportionality to the peak density of the first water layer. Recently, Ma et al. further demonstrated from the NEMD simulations that diagonal charge decoration on the interfacial graphene sheets can substantially reduce the Kapitza resistance between graphene and water by up to 97%, compared to the case without charge decoration. Moreover, the interfacial thermal resistance between few-layer graphene and Si/Ge interface, and corresponding external effects, such as strain, intercalation and interface roughness, are also extensively investigated by NEMD method. In addition, the interfacial thermal resistance can also be used to interpret the interfacial interaction and scatterings or construct coherent phonon transport channel.

To understand the underlying mechanism of phonon scattering at the interfaces, Chalopin et al. proposed a microscopic approach for estimating the spectral phonon transmission across a solid-solid interface based on EMD simulations. In addition, Sääskilähti et al. distinguished the effect of elastic and first-order inelastic phonon scattering on the interfacial thermal conductance through NEMD simulations, and then extended to calculate the frequency dependent phonon mean free paths. Moreover, Zhou and Hu further improved this method by considering all possible three-phonon scattering channels at the interface. Actually, the full order of phonon scatterings can be included by directly using the interatomic force between two atoms at interface as discussed in Ref. 167. Furthermore, Gordiz and Henry developed a new method, based on EMD simulations with full inclusion of anharmonicity, for evaluating the modal contributions to interfacial thermal conductance.

### 4.2 Analysis methods

**Density of states**

Observation of vibrational modes from density of states, known as vibrational density of states (vDOS), is a powerful tool and simple method to characterize the phonon activities in materials. Moreover, density of states is sensitive to many perturbations, such as the temperature dependent phonon scatterings, strain, doping and defects, and so on. The density of states spectra can be simply extracted from MD runs, via the Fourier transform of the velocity autocorrelation function as

$$ vDOS(\omega) = \frac{1}{N\sqrt{2\pi}} \int e^{i\omega t} \left\langle \sum_{j=1}^{N} \mathbf{v}_j(t) \cdot \mathbf{v}_j(0) \right\rangle dt, \tag{62} $$

![Fig. 12](image-url) **Fig. 12** Analysis methods for the thermal conductivity modeling in MD simulations. (a) The suppression of precipitation ratio from Si nanowires (SiNWs) to Si nanotubes (SiNTs). Reproduced with permission from Ref. 170. Copyright 2010 American Chemical Society. (b) Lorentz function fitting in spectral energy density (SED) analysis for the calculations of phonon relaxation time. Reproduced from Ref. 171, with permission from American Physical Society.
where \( \mathbf{v}(t) \) is the velocity vector for particle \( j \) at time \( t \), \( \omega \) is the vibrational frequency, and \( N \) is the number of atoms in the system.

The calculated vDOS has distribution of peaks, which can provide information about the phonon activities from the peak frequency and peak height. For example, in the study of the interfacial thermal transport behaviors in few-layer graphene, the cross-plane strain induced a frequency shifting of the zone-center phonons (less than 4 THz), leading to the increasing of interfacial thermal resistance.\(^{126}\) On the other hand, the suppressed intensity (peak height) of vDOS is often used to explain the reduced thermal conductivity, such as in hydrogenated graphene sheets,\(^{107}\) graphene hydride materials,\(^{162}\) etc. Moreover, the mismatch of vDOS on the two side of an asymmetric nanostructure is often used to interpret the thermal rectification phenomena, corresponding to the asymmetric phonon activities in two opposite thermal transport directions.

### Participating ratio

The participating ratio is another effective way to provide insight on the phonon activities, especially for phonon localization phenomena. Phonon localization is a common phenomenon in thermal transport, which can be induced by the atomic defects, interfaces, edges/surfaces and structure deformation.\(^{177}\) It can serve as a criterion to study the phonon localization degree and its effects on heat conduction. From the lattice dynamics, the phonon participating ratio \( P_\alpha \) of each eigen-mode \( \lambda \) can be quantitatively defined as\(^{176}\)

\[
P_\alpha = \frac{1}{N} \frac{\sum_\lambda |\lambda_{\alpha,\lambda}|^2}{\sum_\lambda |\lambda_{\alpha,\lambda}|^2}, \tag{63}
\]

where \( N \) is the total number of atoms, and \( \lambda_{\alpha,\lambda} \) is the \( \alpha \)th eigenvector component of eigen-mode \( \lambda \) for the \( \alpha \)th atom. \( P_\alpha \) is a dimensionless quantity ranging from 0 to 1, with \( \sim 0 \) denoting the fully localized mode. Here, the eigenvector \( \mathbf{e} \) can be obtained from the lattice dynamic calculations by solving the dynamic matrix of the studied system. The suppression of the \( P_\alpha \) by structure engineering is usually regarded as the signals for inducing phonon localization, as shown in Fig. 12, further leading to the reduction of thermal conductivity.

Besides the frequency perspective, the participating ratio can also be converted to provide a spatial distribution of phonon energy,\(^{125}\)

\[
E_{\alpha} = \sum_\alpha \left(n + \frac{1}{2}\right) \hbar \omega D_\alpha(\omega), \tag{64}
\]

Where \( n \) is the phonon occupation number given by the Bose-Einstein distribution, and \( D_\alpha(\omega) = \sum_{\lambda,\lambda'} \delta (\omega - \omega_{\lambda,\lambda'}) \) is the phonon density of states on atom \( i \) at frequency \( \omega \). One can selectively include propagating or localized phonons into the spatial energy distribution in Eq. (64.) based on the participating ratio of each phonon mode.\(^{125}\)

The participating ratio and spatial energy distribution are widely used to reveal the phonon confinement and thermal rectification in asymmetric nanostructures.\(^{173,177,178}\) For example, Zhang et al.\(^{173}\) has used this method to witness the phonon localization in the surface of silicon nanocone and understand its unique thermal rectification transition with asymmetric degree.

The spectral participation ratio \( P(\omega) \) can also be calculated directly from the MD simulations at arbitrary temperature, without lattice dynamic calculation. Without the polarization information, the participating ratio \( P(\omega) \) is defined as\(^{175}\)

\[
P(\omega) = \frac{1}{N} \frac{\sum_\lambda vDOS(\omega)^2}{\sum_\lambda vDOS(\omega)^2}, \tag{65}
\]

in which the vDOS(\(\omega\)) is the local density of states of \(\alpha\)th atom calculated from the Fourier transform of the velocity autocorrelation function based on Eq. (62.). Different from the Eq. (63.), the calculated participating ratio \( P(\omega) \) by Eq. (65.) can include the all-order of anharmonic scatterings implicitly from MD simulations.

### Normal mode analysis

The anharmonic lattice dynamics calculations are currently limited to three-phonon scattering or four-phonon scattering events due to the computational complexity. Thus, its application is limited to low temperature at which the high order anharmonic interaction is relatively weak. In contrast, the phonon spectral energy density (SED) can be calculated directly from MD simulation, in which the full anharmonicity of the interatomic interactions and other phonon scatterings are simultaneously involved.\(^{173,172,170,181}\)

The SED is usually refereed as \( \Phi(\mathbf{k},\omega) \), where \( \mathbf{k} \) is the wavevector and \( \omega \) is the frequency. For a given phonon mode with wavevector \( \mathbf{k} \) and branch index \( \lambda \), the normal mode coordinate \( \hat{q}(\mathbf{k},\lambda) \) can be obtained by following equation,

\[
\hat{q}(\mathbf{k},\lambda) = \sum_\gamma \frac{m_\gamma}{N} \mathbf{x}_\gamma(t) \cdot \mathbf{e}_\gamma(\mathbf{k},\lambda) \cdot \exp \left[ i \mathbf{k} \cdot \mathbf{x}_\gamma(t) \right], \tag{66}
\]

where \( m_\gamma \) is the mass of the \( \gamma \)th atom in the unit cell, \( \mathbf{x}_\gamma(t) \) and \( \mathbf{e}_\gamma(\mathbf{k},\lambda) \) is, respectively, the position and velocity of the \( \gamma \)th atom at time \( t \), and \( \mathbf{e}_\gamma(\mathbf{k},\lambda) \) is complex conjugate of the eigenvector.\(^{180}\) The SED \( \Phi(\mathbf{k},\omega) \) can be calculated by the following formula,

\[
\Phi(\mathbf{k},\omega) = \sum_\lambda \left| \hat{q}(\mathbf{k},\lambda) \right|^2 \exp(-i\omega t) dt, \tag{67}
\]

The previous study\(^{180}\) found that the eigenvectors in Eq. (66.) is a necessary part in the normal mode analysis to accurately predict the phonon relaxation time and mean free path. Without the eigenvectors, only the phonon dispersion curves can be accurately predicted.\(^{170,181}\) On the other hand, Feng et al.\(^{182}\) provided both analytical and numerical proof that the eigenvectors are not absolutely necessary in SED calculations, and two different definitions of SED give the same result within the numerical accuracy. However, if the frequencies of phonon modes are too close to each other, the eigenvectors are necessary in order to separate them into individual parts.\(^{182}\)

In addition, the phonon relaxation time \( \tau(\mathbf{k},\omega) \) can be obtained by fitting the SED curve by Lorentzian function,\(^{173,180}\)

\[
\Phi(\mathbf{k},\omega) = \frac{I}{1 + \left[\frac{\omega - \omega_0}{\gamma}\right]^2}, \tag{68}
\]

where \( I \) is the peak intensity, \( \omega_0 \) is the frequency of peak center, and \( \gamma \) is the half-width at half-maximum, as shown in Fig. 12. Finally, the \( \tau \) at each wavevector \( \mathbf{k} \) and frequency \( \omega \) is defined as:

\[
\tau(\mathbf{k},\omega) = \frac{1}{2\gamma}, \tag{69}
\]

With the information of spectral \( \tau \) from SED, the frequency contribution to \( \kappa \) can be evaluated from the kinetic theory,

\[
\kappa(\omega) = C_T (\omega) v(\mathbf{k},\omega) \tau(\mathbf{k},\omega), \tag{70}
\]

where \( C_T(\omega) \) is the phonon volumetric specific heat at frequency \( \omega \), \( v(\mathbf{k},\omega) \) is the phonon group velocity at wavevector \( \mathbf{k} \) and frequency \( \omega \). As shown in Eq. (66.), only the positions and momenta of the atoms from MD simulations are evolved in SED analysis. Therefore,

© Engineered Science Publisher LLC 2018

ES Energy & Environment, 2018, 1, 16–55 | 31
it is also an effective way to provide insight to the external effects on phonon scattered, phonon frequency and thermal transport properties.

For example, using SED analysis, Qiu et al. found that the presence of silicon dioxide substrate induced substantial reduction of phonon relaxation time compared to that in suspended single-layer graphene. Zhang et al. demonstrated that bulk hexagonal boron nitride is a more appealing substrate to achieve high performance heat dissipation in supported graphene, corresponding to a less affected phonon relaxation time and dominant phonon mean free path as applying SED analysis. Moreover, the SED method is also used to interpret the suppression of thermal conductivity in defected and doped graphene. Hu et al. found that phonon modes with long mean free path are strongly suppressed, resulting in the suppressed size dependence and the weaker temperature dependence of the thermal conductivity compared to the pristine graphene.

**Mode decomposition**

As spectral phonon properties can be obtained from the normal mode analysis and SED analysis in EMD simulations, relevant methods are developed to extract frequency/mode-dependent thermal transport information in the framework of NEMD. Zhou et al. proposed the time domain and frequency domain direct mode decomposition method (TDDDD and FDDDD) by calculating the spectral/modal heat current in NEMD simulations. The TDDDD method is based on the correlation function of atomic force and velocity as

\[ K_p(t) = \frac{1}{2} \langle F_p(t_1) \cdot [v_i(t_1) + v_j(t_2)] \rangle, \]

and the spectrum of atomistic heat current is related to the Fourier transform of Eq. (66).

\[ J_p(\omega) = \text{Re} \int_{-\infty}^{\infty} e^{i\omega \tau} K_p(\tau) d\tau. \]

In the light of the heat current expression in MD, the spectral heat current in a control volume can be obtained as

\[ J(\omega) = \frac{1}{\hbar} \sum_i \sum_j \sum_{S_p} \sum_{v_p} \sum_{S_e} \sum_{v_e} \sum_{S} \sum_{v} J_p(\omega) \cdot (r_p^i - r_p^j) \]  \hspace{1cm} (73)

The individual phonon contributions to overall thermal conductivity are defined as \( \kappa(\omega) = \frac{1}{\rho \pi T} J(\omega) \). As for the TDDDM method, the heat current of each mode can be defined as

\[ J_p = \sum_j \left< \frac{1}{\sqrt{Nm_i}} [E_p, e_j] \exp(\mathbf{k} \cdot \mathbf{r}_p) \hat{X}_j - S_p, e_j \exp(\mathbf{k} \cdot \mathbf{r}_p) \hat{X}_j \right> \]  \hspace{1cm} (74)

Furthermore, the phonon mode contributions to the overall thermal conductivity are derived from the formula, \( \kappa = \frac{1}{\rho \pi T} J_p \). The TDDDM method can resolve the thermal conductivity into mode contributions while the FDDDD method is only available for spectral analysis. Both methods can be used to investigate the size effect of thermal conductivity from the aspect of phonon contributions, which cannot be solved by SED or normal mode analysis within the EMD framework.

Interfacial thermal transport plays an essential role in nanoscale structures and composite materials. The phonon scattering at the interface results in the Kaptiza resistance, which can be simply described by the acoustic mismatch model (AMM) and the diffusive mismatch model (DMM). In addition, the Green’s function method provides a more precise depiction of heat conduction at the interface. However, these methods are unable to directly account for the inelastic effects due to the interatomic anharmonic forces. To assess the importance of inelastic scattering at planar interfaces, Sääskilähi et al. developed a method to separate the elastic and inelastic contributions to interfacial thermal...
conductance. With the expansion of interatomic forces, the correlation function of the interface can be written as

\[ K_i(\tau) = \frac{1}{2} \sum_{\beta} k_{\beta}^0 A_\beta^i(\tau) + \frac{1}{4} \sum_{\beta, \gamma} \gamma_{\beta \gamma}^i B_\gamma^i(0, \tau) + \cdots, \tag{75} \]

where only the first two terms are given. Based on the Fourier's transform of Eq. (75.), the spectral decomposition of the conductance is resolved into elastic and inelastic contributions as

\[ g^{\text{elastic}}(\omega) = \frac{1}{A T} \sum_{\alpha} \sum_{\beta} k_{\beta}^0 \Re \{ \hat{X}_{\alpha}(\omega) \}, \tag{76} \]

\[ g^{\text{inelastic}}(\omega, \omega') = -\frac{1}{A T} \sum_{\alpha, \beta} \gamma_{\alpha \beta}^{\text{inelastic}} \Re \{ \hat{X}_{\beta}(\omega) \}, \tag{77} \]

\[ G = \frac{1}{A T} \int_0^{2\pi} \frac{[g^{\text{elastic}}(\omega) + g^{\text{inelastic}}(\omega)^\ast]}{2\pi} d\omega. \tag{78} \]

The accuracy of \( g^{\text{inelastic}}(\omega) \) is determined by the higher-order terms of Eq. (75.) as the truncation of the interfacial phonon force expansion is unavoidable. Since this method can provide the spectral thermal conductance at the interface, we can obtain the dominant phonon frequencies of the heat conduction as well as the temperature-dependent elastic and inelastic contributions. It was found that inelastic effects strongly facilitate energy transfer across the interface at high temperatures, and the harmonic methods like Green's function are ineffective to explain such effects where anharmonic interactions govern the interfacial thermal transport.

The above decomposition methods from Refs. 16-18 and mainly focus on the calculation of spectral/modal heat current, and determine the spectral/modal thermal conductivity through the Fourier's law assuming the same temperature gradient for all phonon modes. However, the local thermal equilibrium assumption usually breaks down at nano/micro-scale. Feng et al. developed the spectral phonon temperature (SPT) method to calculated the spectral/modal thermal conductivity through the Green's function are ineffective to explain such effects where anharmonic interactions govern the interfacial thermal transport.

The SPT method is able to illustrate the mode temperature distribution of ballistic and diffusive phonons, so it is very useful to investigate local thermal non-equilibrium phenomena in nanomaterials and across interfaces. Moreover, the thermal conductivity can be decomposed into mode contributions on the basis of TDDDM method and SPT method within the framework of NEMD.

\[ T_j = \langle \hat{X}_j^0(\tau) \hat{X}_j^0(\tau) \rangle, \tag{79} \]

The excited wave packet usually has very small amplitude \( A \), such as 0.01 Å, to justify the use of the harmonic approximation. By Fourier transform of the wave-packet, we can get a single peak in k-space centered at \( K \), with the full-width at half-maximum (FWHM) of \( 2\pi/\eta \). The transport behaviors of wave-packet can be visualized by recorded the time-dependent of wave packet propagating (See Fig. 13). Moreover, by computing the total energy of incident and transmitted wave-packet, one can obtain the energy-transmission coefficient, which is helpful for analyzing the thermal transport ability through the interface and surface. It should be noted that the wave-packet method is usually performed at low temperature, at which the anharmonic phonon-phonon scatterings can be ignored. At high temperature, it is difficult to distinguish the wave packet signal from the temperature noise, and the wave packet is easily scattered by the atom vibration at high temperature.

By performing wave packet simulations, Chen et al. elucidate that in the coherently bonded graphene–carbon nanotube hybrid structure, the transmission coefficient across graphene sheets for various phonon modes can be significantly promoted by the covalent bonding compared to that with non-bonded weak inter-layer interaction, leading to the remarkably improved c-axis thermal transport in the hybrid structure for high-performance thermal dissipation. From wave packet simulations, Shao et al. demonstrated that phonon-boundary scatterings strongly dependent on the boundary conditions.

4.3 Other MD based methods

Two-temperature model MD: All the above MD methods for the calculation of thermal conductivity have only accounted for phonons as the dominant heat carrier, while ignoring the contribution of electrons. When the contributions from electron or electron-phonon interaction to heat conduction become important, such as in metals or semimetals, other more reasonable method should be proposed.

The two-temperature model (TTM) is the approach that can properly include electron-phonon coupling, in which the electron and phonon are treated as two separated subsystems and a
temperature is assigned to each subsystem, as shown in Fig. 13. Compared to the classical MD simulations, more energy transport channels such as electronic and electron-phonon coupling channel are added, leading to better approximation to realistic energy transport events in electron dominant materials. Wang et al. successfully applied this TTM-MD method to capture the non-equilibrium between electron and phonon in thermal transport across metal-nonmetal interfaces. The simulation results are physically sound and are in better agreement with experimental data compared to those obtained using conventional MD simulations.

**Approach to equilibrium MD:** The approach to equilibrium MD, AEMD, is another kind of thermal conduction modeling method. The studied system is initially set out-of-equilibrium by partly heated system portion at a different temperature from the rest. Then, when the system approaches to equilibrium, the time evolution of the physical properties, such as temperature difference and total energy between the two parts, are monitored. Lampin et al. demonstrated that for most practical cases of interest, the temperature decay is exponential and can be used to extract the thermal conductivity of homogeneous materials. Compared to the expensive computation of autocorrelation function in EMD and establishing a non-equilibrium steady state in NEMD, the equilibrium in AEMD can be reached in a few tenths to hundreds of ps with much reduced computational cost.

### 5. Non-equilibrium Green's function

Despite the popularity of BTE-based methods and MD simulations to model phonon transport, these methods that have been discussed in the previous sections suffer from some shortcomings when applied to phonon quasi-ballistic transport, which usually happens at low temperature or in the devices with small size. In the BTE calculations, lattice vibration is treated as incoherent particles, i.e. phonons, which is a good approximation when the phonon wavelength is much smaller than the characteristic size of the nanostructures. However, when the characteristic size is comparable to or larger than the phonon wavelength, the lattice vibrations would exhibit wave features, such as interference, diffraction and localization, which cannot be captured by the picture of incoherent particle. As an alternative, lattice dynamics approach models phonons as coherent waves. But, it is difficult to apply this approach to the systems with complicated atomic structures, such as amorphous layers and the crystal with defects or multiple interfaces, since lattice dynamics is more suitable for periodic lattice systems. Instead of treating lattice vibrations as particles or waves, MD simulations do not need to assume the nature of phonon transport. However, the systems in MD simulations are classical so that phonon transport properties at low temperatures is not accurate.

Non-equilibrium Green's function (NEGF) approach, also called atomistic Green's function approach in some literatures, is another powerful tool to study phonon transport in nanostructures. Compared with BTE, MD and lattice dynamics, NEGF, which is based on a dynamical equation and the quantum mechanical distribution for phonons, models lattice vibrations as waves and includes atomic details into account. When the anharmonic interaction is not considered, i.e., phonon-phonon scatterings are absent, the phonon transmission and thermal conductance obtained from the NEGF calculations are exact. Therefore, the NEGF is an ideal approach to investigate the phonon transport in nanostructures where elastic scatterings are dominant. The NEGF approach has been successfully employed to study phonon transport across low-dimensional systems, interfaces and superlattice structures, and so on.

In Sec. 5.1, the general framework of the standard NEGF, the numerical procedures and the methods to generate the inputs for NEGF will be introduced. A few applications of the NEGF method will be presented in Sec. 5.2. In Sec. 5.3, the extension of the NEGF method to calculate the mode-specific phonon transmission will be briefly mentioned.

#### 5.1 Non-equilibrium Green's function approach for phonon transmission

**General Framework**

The typical system in the NEGF approach is illustrated in Fig. 14 (a). The simulation system consists of three regions, one device (scattering) region and two contacts (thermal reservoirs). The temperatures of the two contacts are kept at \( T + \Delta T / 2 \) and \( T - \Delta T / 2 \), respectively, with an infinitely small temperature difference \( \Delta T \). The task of the NEGF is to obtain the frequency-dependent phonon transmission, or the number of phonons passing the device region from one contact to the other one, under the driving of the temperature difference. Using the information of the phonon transmission, the heat flux across the device region could be calculated through the well-known Landauer formalism.

![Fig. 14 (a) Schematic diagram of the simulation system in a typical NEGF calculation, (b) Illustration of the sublayers in the left contact.](image)

The NEGF approach is originated from the dynamical equation for lattice vibrations. Under the harmonic approximation the phonon waves in a closed system can be described by the dynamical equation

\[
(\omega^2 I - H)\Phi(\omega) = 0, \tag{83}
\]

where \( \omega \) is the angular frequency of lattice vibration, \( I \) is the identity matrix, \( H \) is the harmonic matrix. The bold letters present matrices and vectors. The elements \( H_{ij} \) of the harmonic matrix \( H \) is defined as

\[
H_{ij} = \left. \frac{1}{\sqrt{M_i M_j}} \frac{\partial^2 U}{\partial u_i \partial u_j} \right|_0, \tag{84}
\]

where \( u_i \) and \( u_j \) denote the i-th and j-th atomic displacement degrees of freedom in the system; \( M_i \) and \( M_j \) are the atomic masses of the atoms that the i-th and j-th degrees of freedom belong to; \( U \) is the interatomic potential energy of the system; \( |_0 \) means that the derivative is calculated when atoms stay in their equilibrium positions, i.e., \( u_i = u_j = 0 \). \( \Phi(\omega) \) is the eigenvector of \( H \), and could be interpreted the magnitude of the vibrational modes, in which the m-
th element, $\Phi_m$ is the product of $i \omega$ and $\sqrt{\Sigma_m}$. 

Instead of solving Eq. (83.) directly for the phonon waves as in linear lattice dynamics simulations, the Green's function method can be used to obtain the dynamical response of the lattice system under small perturbations, such as small displacements or small forces acting on the atoms. The Green's function corresponding to Eq. (83.) is defined through 

$$(\omega^2 I - H) G = I,$$  

(85)

where $G$ is the Green's function. The element $G_{ij}$ in the Green's function $G$ represents the response of the degree of freedom $j$ to the perturbation on the vibration of the $i$-th degree of freedom.

When a system is kept at a constant temperature through an external heat source or sink, the system becomes an open system and Eq. (83.) has to modified to the form 

$$(\omega^2 + \delta i) I - H G(\omega) = s,$$  

(86)

A small imaginary number $\delta i$ with $\delta \rightarrow 0$ is included in the dynamical equation to represent the damping in the open system. $s$ is a source term that is interpreted as the inflow of phonons into the system. Similar to Eq. (85.), the Green’s function of the open system is defined as $g = [(\omega^2 + \delta i) I - H]^{-1}$.

With the above concepts of Green’s function, the phonon transmission through a device or a scattering region, which could be a nanostructure, an interface, etc., could be calculated. Based on the three-region divisions of the simulation system, as illustrated in Fig. 14a, the harmonic matrix of the entire system could be written as 

$$H = \begin{bmatrix} H_{L,L} & H_{L,D} & 0 \\ H_{L,D} & H_{D,D} & H_{D,R} \\ 0 & H_{R,D} & H_{R,R} \end{bmatrix},$$  

(87)

where the subscripts $L$, $R$ and $D$ represent the left contact, the right contact and the device region; the submatrices $H_{L,L}$, $H_{L,D}$, $H_{D,D}$, and $H_{D,R}$ represent the connection matrices whose elements involve the harmonic constants corresponding to one atom in the device region and the other one in the contacts. The matrices indicate that any atoms in one contact do not interact with the atoms in the other contact. When the two contacts are connected to the external heat source or heat sink, the dynamical equation of the entire system can be expressed as 

$$\begin{bmatrix} (\omega^2 + \delta i) I - H \end{bmatrix} \begin{bmatrix} \Phi_L^i + \chi_L^i \\ \Phi_D^i \\ \Phi_R^i + \chi_R^i \end{bmatrix} = \begin{bmatrix} s_L \\ 0 \\ s_R \end{bmatrix},$$  

(88)

where $\Phi_L^i$ and $\Phi_R^i$ are the eigenvectors of the harmonic matrices of the isolated contacts; $\Phi_D^i$ are that of the device that are coupled with two contacts; $\chi_L^i$ and $\chi_R^i$ represent the change of atomic displacements of the two contacts due to the coupling with the device region; $s_L$ and $s_R$ are the source terms for the two contacts due to energy transfer with the external surrounding. From the first row of Eq. (88.), we have

$$\begin{bmatrix} (\omega^2 + \delta i) I - H_{L,L} \end{bmatrix} \Phi_L^i + \left[ (\omega^2 + \delta i) I - H_{L,L} \right] \chi_L = \left[ (\omega^2 + \delta i) I - H_{L,D} \right] \Phi_D = s_L.$$  

(89)

Since the left contact has energy transfer with the external thermostat, according to Eq. (86.) the first term of Eq. (89.) equals $s_L$. Thus, Eq. (89.) becomes $\chi_L = g_L H_{L,D} \Phi_D$. Similarly, the third row of Eq. (88.) leads to $\chi_R = g_R H_{R,D} \Phi_D$. $g_L$ and $g_R$ are the Green's functions of the two contacts, which are also called retarded surface Green's functions and have the form 

$$g_L = \left[ (\omega^2 + \delta i) I - H_{L,L} \right]^{-1},$$  

(90)

$$g_R = \left[ (\omega^2 + \delta i) I - H_{L,L} \right]^{-1}.$$  

(91)

Based on the obtained expressions for $\chi_L$ and $\chi_R$, the eigenvectors for the coupled contacts are

$$\Phi_L = \Phi_L^i + \chi_L = \Phi_L^i + g_L H_{L,D} \Phi_D,$$  

(92)

$$\Phi_D = \Phi_D^i + \chi_R = \Phi_D^i + g_R H_{R,D} \Phi_D.$$  

(93)

Furthermore, substituting $\chi_L$ and $\chi_R$ into the second row of Eq. (88.), the atomic displacement vector for the device region is then written as

$$\Phi_D = G_{33} (H_{DD} \Phi_D^i + H_{DR} \Phi_R^i),$$  

(94)

where $G_{33} = (\omega^2 I - H_{DD} - \Sigma_D - \Sigma_R)^{-1}$ is the Green's function for the device region with the self-energy matrices $\Sigma_D = H_{DD} \Phi_D \Phi_D^i$ and $\Sigma_R = H_{DR} \Phi_R \Phi_R^i$.

From Eq. (92.) to Eq. (94.), the wave functions of both the device region and the contacts have been obtained, which are linked with the Green's functions of the three regions. Using the information of the obtained wave functions, the phonon propagation could be determined as the following. The total energy of the system can be decomposed into the atomic energies,

$$E = \sum_p E_p,$$  

(95)

with 

$$E_p = \frac{1}{2} M_p \dot{u}_p \cdot \dot{u}_p + \frac{1}{4} \sum_q k_{pq}^2 \dot{u}_q \cdot \dot{u}_q,$$  

(96)

where $E_p$ is the energy associated with the $p$-th degree of freedom, and $k_{pq} = H_{pq} \sqrt{M_p M_q}$. The first term of Eq. (96.) is the kinetic energy, while the second term is the potential energy. Using Newton's second law, $M_p \ddot{u}_p = -\sum_q k_{pq} u_q$, the time derivative of $E_p$ is rewritten as

$$\frac{dE_p}{dt} = \frac{1}{4} \sum_q \left( u_q^2 k_{pq} u_p + \dot{u}_q^2 k_{pq} u_p - \dot{u}_q^2 k_{pq} u_p - \ddot{u}_q^2 k_{pq} u_p \right).$$  

(97)

For a given phonon, the atomic displacement for the $p$-th degree of freedom is connected with the wave function of the phonon and expressed as $u_p = \Phi_p \exp(-i \omega t) / \sqrt{M_p}$. Substituting the expression of atomic displacement into Eq. (97.), the time derivative of $E_p$ is now associated with the eigenvectors of the vibrations, and Eq. (97.) can be further simplified as.
\[ \frac{dE_d}{dt} = \frac{\alpha}{2\Gamma} \sum_{\ell} (\Phi_\ell^d \Phi_\ell \Phi_\ell^d - \Phi_\ell^d \Phi_\ell \Phi_\ell). \] (98)

By decomposing \( dE_d/\Gamma \) into the contributions from different degrees of freedom, the local heat current between the degree of freedom \( p \) and \( q \) can be defined as
\[ J_{pq} = \frac{\alpha}{2\Gamma} (\Phi_\ell^d \Phi_\ell \Phi_\ell^d - \Phi_\ell \Phi_\ell^d \Phi_\ell). \] (99)

Summing up the local heat current and normalizing the energy of the phonons at frequency \( \omega \) to \( h\omega \), the heat current due to a phonon from the device region to the right contact can be derived as
\[ J(\omega) = \frac{h}{2} \text{Tr} (\Phi_\ell^d \Phi_\ell \Phi_\ell^d - \Phi_\ell \Phi_\ell^d \Phi_\ell). \] (100)

Substituting Eq. (92.) and Eq. (93.) in Eq. (100.), the heat current for one particular phonon frequency is related to the Green's functions through
\[ J(\omega) = \frac{h}{2\Gamma} \text{Tr} (\Gamma_\ell \Gamma_\ell^* \Gamma_\ell \Gamma_\ell^*). \] (101)

where \( \Gamma_\ell = i(\Sigma_\ell - \Sigma_\ell^*) \), \( \Gamma_\ell^* = i(\Sigma_\ell - \Sigma_\ell^*) \) and \( \ast \) denotes the conjugate transpose of the matrix. Using the Landauer formalism, the thermal conductance through the device region from the left contact to the right contact is derived as
\[ \sigma = \frac{J}{A} = \frac{1}{2\pi A} \int d\omega \omega \frac{dJ(\omega)}{dT} \text{Tr} (\Gamma_\ell \Gamma_\ell^* \Gamma_\ell \Gamma_\ell^*). \] (102)

The total phonon transmission across the device region is defined as
\[ \Xi(\omega) = \text{Tr} (\Gamma_\ell \Gamma_\ell^* \Gamma_\ell \Gamma_\ell^*). \] (103)

In addition to the total phonon transmission to count the number of phonons traveling through the device region, the phonon transmittance, \( \xi(\omega) \), is also used in some literatures to present the phonon transport process, which is linked with \( \Xi \) through
\[ \Xi(\omega) = M(\omega) \xi(\omega). \] (104)

where \( M(\omega) \) is the total number of phonon modes at frequency \( \omega \) from the left contact. A simulation system made up of a pure material can be employed to calculate \( M(\omega) \). In such calculations, \( \xi(\omega) \) equals 1 so that \( M(\omega) \) is simply identical to the phonon transmission in the pure material according to Eq. (104.).

**Self-energies**

As indicated in Eq. (94.), the crucial step to calculate the phonon transmission is to compute the self-energies \( \Sigma_\ell \) and \( \Sigma_\ell^* \) of the two contacts. The challenge of calculating the self-energies is that the contacts are semi-infinite, thus the dimensions of the matrices \( \Sigma_\ell \) and \( \Sigma_\ell^* \) are large, leading to the difficulty in computing the surface Green's functions \( \Phi \) and \( \Phi^* \).

Here, we take the left contact as an example to show the procedures to obtain the surface Green's functions. The surface Green's function for the right contact can be derived in the same way. Since in the Green's function calculations only \( \Phi_{10} \) and \( \Phi^*_{10} \) are required, where \( \Phi_{10} \) and \( \Phi^*_{10} \) have only a small portion of non-zero elements, which represent the interaction between the device region and the contact, it is not necessarily to solve the entire \( \Phi \) matrix, but a submatrix in \( \Phi \). Considering that the interatomic interaction is short-ranged, the left contact can be divided into a few sublayers, as illustrated in Fig. 14b. Due to the semi-infinite nature of the contact, these sublayers are assumed to identical to each other. Since the atoms in a layer only interact with those in the neighboring layers, the matrix \( \Phi_{10} = (\delta^2 + \delta^0) \mathbf{1} - \Phi_{11} \) can be expressed as
\[ \Phi_{10} = \begin{bmatrix} M_{11} & M_{10} \end{bmatrix}, \] (105)

where \( M_{11} \) is the dynamical matrix for the \( i \)-th sublayer in the left contact; and \( M_{10} \) (\( M_{01} \)) represents the coupling between the \( i \)-th and \((i+1)\)-th sublayers. Taking the advantage of the translational symmetry of the sublayers, we have \( M_{1i} = M_{10} \) and \( M_{0i} = M_{1i} \)

\[(i = 1, 2, ...). \]

The corresponding Green's function \( g \) can be similarly decomposed to a series of submatrices, denoted as \( g^{\alpha} \). According to the definition of \( g \), the submatrices have to obey the following relations
\[ g^{\alpha}_{10} g^{\alpha}_{01} - M_{10}^{\alpha} M_{01}^{\alpha} = 1, \] (106)
\[ -M_{10}^{\alpha} g^{\alpha}_{01} + M_{01}^{\alpha} g^{\alpha}_{01} - M_{10}^{\alpha} M_{01}^{\alpha} = 0, \] (107)

Many methods have been developed to obtain \( g^{\alpha} \) from Eq. (106.) to Eq. (107.). The details of these methods could be found in Ref. 218.

One commonly used approach is called decimation method. The basic idea is to eliminate \( g^{\alpha}_{m-n,0} \) (\( m \) is a nonnegative integer) by expressing it using \( g^{\alpha}_{m-1,0} \) and \( g^{\alpha}_{m,1} \) through Eq. (107.). Inserting the expression of \( g^{\alpha}_{m-1,0} \) into Eq. (106.) and Eq. (107.), the set of equations become in terms of \( \{ g^{\alpha}_{m,1}, g^{\alpha}_{m-1,0}, g^{\alpha}_{m,0}, g^{\alpha}_{m,2,0}, \ldots \} \). Since the new equations have the same form as Eq. (106.) and Eq. (107.), we can perform the above procedures to eliminate \( \{ g^{\alpha}_{m,1}, g^{\alpha}_{m-1,0}, g^{\alpha}_{m,0}, g^{\alpha}_{m,2,0}, \ldots \} \) and obtain the equations with respect to \( \{ g^{\alpha}_{m,1}, g^{\alpha}_{m,0}, g^{\alpha}_{m,2,0}, \ldots \} \). Repeating such elimination process \( n \) times, we reach the following relation
\[ M_{10}^{\alpha} g_{01}^{\alpha} - M_{10}^{\alpha} g_{01}^{\alpha} = 1, \] (108)
\[ -M_{10}^{\alpha} g_{01}^{\alpha} + M_{01}^{\alpha} g_{01}^{\alpha} - M_{10}^{\alpha} M_{01}^{\alpha} = 0, \] (109)

where \( M_{10}^{\alpha}, M_{10}^{\alpha}, \) and \( M_{01}^{\alpha} \) are the coefficient matrices generated in the elimination process, which can be generated in an iterative way starting from \( M_{10}^{\alpha}, M_{10}^{\alpha}, \) and \( M_{01}^{\alpha} \). It can be easily proved that as \( n \) becomes sufficiently large, \( M_{10}^{\alpha} \rightarrow 0 \) and \( g_{01}^{\alpha} = (M_{01}^{\alpha})^{-1} \). Physically, this is because the perturbation on the sublayer far away the device region should have negligible effects on 0-th sublayer.

**Interatomic force constants**

According to the discussions above, the only required inputs for NEGF method are the harmonic force constants of the system studied. Often the empirical interatomic potentials are employed in
NEG calculations to describe interatomic interactions. However, the accuracy of the existing empirical potentials is usually not satisfactory to fully reproduce the vibrational properties of the materials. Therefore, quite a few attempts have been devoted to integrate first-principles calculations with NEG calculations.

Comparing to extracting first-principles interatomic force constants for calculating thermal conductivity of bulk crystals, which has been discussed in Sec 3.2, the process becomes computationally challenging when applied to the systems that are often explored by the NEG method. This is because the periodicity is usually lost in these systems, where the interfaces or defects usually exists, and large supercells would be required for the first-principles calculations to capture the essential characteristics in the system investigated. A few solutions have been proposal to boost the efficiency of extracting harmonic force constants from first-principles.

To relieve the computational burden of modeling the crystalline systems made up of two different atomic species with similar chemical properties, the mass approximation that was employed to calculate the thermal conductivity of alloys has thus been used to extract interatomic force constants from the first principles. Under the mass approximation, the differences in the lattice constants and the force fields between the two dissimilar materials are ignored and only the difference in atomic mass is taken into account. Therefore, the harmonic constants can be extracted in a smaller supercell using the small-displacement method or the primitive unit cell through the perturbation approach by just following the procedures discussed in Sec. 3.2 for bulk crystal. Despite the simplicity of the mass approximation, the local force-field difference is neglected, which might affect phonon transport considerably.

To overcome the limitations of the mass approximation, Gu et al. employed the higher-order force constant method to extract the harmonic force constants. They first approximated the atomic system of large dimension that contains two species as a virtual crystal. In the virtual crystal, the two types of atoms in the first-principles simulation are replaced by virtual atoms, whose pseudopotential is the percentage-weighted pseudopotentials of the two types of elements through

\[
V_\sigma = \frac{1}{2}V_{\text{elem}} + \frac{1}{2}V_{\text{elem2}},
\]

where \(V_{\text{elem}}\) and \(V_{\text{elem2}}\) are the pseudopotentials for element 1 and element 2 and \(\sigma\) represents the likelihood of the virtual atom to be element 1 or 2. \(\sigma = 1\) and -1 represents elements 1 and 2, respectively. Comparing the realistic atomic system with the virtual crystal, the difference of their total energies originates from two aspects: (1) the atoms are not uniformly distributed as in a virtual crystal but are of small displacements \(u\) from their virtual crystal counterparts. (2) The type of the atom \(\sigma\) and the corresponding force field in the real system differs from the virtual atom. Both of them make the harmonic force constants of the real system deviate from the virtual crystal. The higher-order force constant model approximates the difference between the total energy of the real lattice system and that under the virtual crystal approximation using the Taylor's expansion of the total energy of the reference virtual crystal with respect to \([\sigma a]\) and \([\sigma a]\). These coefficients in the Taylor's expansion reflect the vibrational and chemical properties of the virtual crystal, which can be obtained from first-principles following the steps presented in Ref. 210. Once these coefficients are available, the Taylor's expansion is indeed the expression of the energy of the system and could be used in MD simulations to relax the structure. Furthermore, the harmonic force constants could be easily calculated with these coefficients by taking the second-order derivatives of the energy with respect to atom displacements.

For more complicated systems, such as those involving two dissimilar materials with totally different crystal structure, neither the mass approximation nor higher-order force constant method could be used. Instead of including the whole system in first-principles calculations, in some previous studies the system is divided into a few subsystems, and the harmonic force constants of each subsystem are extracted independently from the first-principles. As the computational burden increases rapidly with the size of the simulations, such an approach can boost the efficiency of extracting harmonic force constants reasonably.

5.2 Applications

Phonon transport across one-dimensional systems

Low-dimensional materials, such as nanotubes, nanowires and twodimensional materials, have attracted tremendous attention due to their unique physical properties for fundamental science and various technological applications. The thermal properties of low-dimensional materials are crucial for many applications, such as thermoelectrics and thermal management. Since phonon transport properties in low-dimensional materials are different from those in three-dimensional bulk materials, the heat conduction in low-dimensional systems were intensive explored in the past two decades. Due to the simplicity of the formulism, the NEG was commonly used as a theoretical tool to study phonon transport in low-dimensional systems. The ballistic limits of thermal conductance of low-dimensional materials and the effects of many physical factors like size, defects, strain on the phonon transmission and thermal conductance are usually explored.

Carbon nanotubes are typical one-dimensional nanostructures. They are believed to possess high thermal conductivity due to their unique phonon band structure and low phonon scatterings. However, the synthesized structures inevitably include some defects, such as vacancies, substitutional defects and chemical defects, which would reduce the thermal conductivity.

Mingo et al. implemented a first-principles method to investigate the transmission of phonons through Stone-Wales defects and nitrogen substitutional impurities. They proposed to use a Lagrange-multiplier technique to impose the translational and rotational invariance of the harmonic force constants. Fig. 15a compared the phonon transmission through a single Stone-Wales defect using the force constants that obey the invariances and that using the untreated force constants. The different transmissions calculated from different sets of force constants for low-frequency phonons emphasize the importance of imposing the invariances. The phonon transmission calculated by the force constants from the Brenner potential was presented in Fig. 15b, and quantitative differences between the results from first-principles and those from the Brenner potential in the positions of peaks, and the span of the spectrum can be clearly observed. In addition, it was found from Fig. 15c that the effect of the nitrogen impurity is rather minor as compared with that produced by structural defects of the Stone-Wales type. When the multiple defects or impurities exist in the nanotubes, the phonon transmission could be quite distinct from the nanotube with single defects. Savic et al. identified that multiple scattering induced interference effects are prominent for isotope
concentrations smaller than 10%; otherwise, the thermal conduction is mainly determined by independent scattering contributions of single isotopes.

Fig. 15 Gray shading shows phonon transmission for (a) Stone-Wales defect from first principles, (b) Stone-Wales defect from Brenner potential, and (c) nitrogen substitutional impurity from first principles, on a (7,0) SWCNT. Black shading shows phonon transmission for the pristine nanotube. Dashed red lines show the unphysical results obtained using unsymmetrized force constants. Reproduced from Ref. 207 with permission from American Physical Society.

Nanoribbons, stripes of two-dimensional materials, which are also quasi-one-dimensional systems, have distinct electronic properties from the infinitely large sheet. The thermal conductance of a few different types of nanoribbons have been studied. The effects of the orientation, width and edge roughness on the phonon transmission in nanoribbons were investigated using the NEGF approach. For instance, the phonon transmission function across graphene nanoribbon was calculated by a few groups, in which the obtained phonon transmissions are quite different for nanoribbons with zigzag edge and armchair edge. With the obtained phonon transmission, the calculated thermal conductance of zigzag nanoribbon is found to be higher than the armchair one with comparable widths. Tan et al. attributed the chirality-dependent thermal conduction to the fact that low frequency bands in graphene nanoribbons with zigzag edges are more dispersive than those in those with armchair edges. MD simulations also confirmed the zigzag nanoribbon is more capable of conducting heat. Another finding from the NEGF study is the normalized thermal conductance (the ratio between thermal conductance and the cross-sectional area of a nanoribbon) exhibits a negative dependence on the width of nanoribbon.

Phonon transmission across interfaces and thermal boundary conductance

Interfaces play a critical role in determining phonon dynamics and thermal conduction in nanostructures. The detailed information of how a phonon quanta with specific energy and momentum is scattered by an interface is essential for designing nanostructures with desirable thermal performance using mesoscopic modeling tools, such as the phonon Boltzmann transport equation (BTE) based method and Monte Carlo simulations. The NEGF method could provide the frequency-dependent phonon transport across interfaces.

Unlike the one-dimensional system discussed above, the interfacial system is usually three-dimensional. Assuming the interface is parallel to the x-y plane, the system is infinitely large in the x and y directions, resulting in infinitely large number of degrees of freedom in the harmonic equation as well as the Green's function. Zhang et al. proposed to use the wave vector representation to construct these matrices. In this approach, the infinitely large system is divided to identical tubular unit cells along the z direction. Due to the periodicity in the x and y directions, the harmonic matrix can be expressed in a plane-wave form in terms of wavevector \( \mathbf{k} \). By averaging the Green's function with respect to wave vector, the mean Green's function can be easily obtained, which is used to determine

Fig. 16 (a) Interface formed between Si and Ge-like materials with an 8% lattice mismatch after relaxation with MD simulation. (b) Frequency-dependent phonon transmittance across the relaxed interfaces formed between Si and Ge-like material with different percentages of lattice mismatch. Reproduced from Ref. 208 with permission from American Physical Society.
phonon transmission and thermal conductance as discussed in Sec. 5.1.

Zhang et al. investigated the phonon transmission across Si/Ge/Si and Ge/Si/Ge interfaces, where the lattice of Si is stretched to match that of Ge. The thermal conductance agrees well with the prediction from the acoustic mismatch model at low temperatures. To consider phonon transmission across realistic interfaces, Li and Yang developed an integrated molecular dynamics (MD) and AGF approach. In their work, the lattices of silicon and germanium crystals that form the interface are not assumed to share the same lattice constant, but have their own lattice constants. They used Tersoff potential to describe the interatomic interaction of these two materials. For a Si/Ge interface, 25 unit cells of Si are included to match the cross section of 24 unit cells of Ge. In order to quantitatively see how the lattice mismatch affects the phonon transmission across the interface, they modified the potential of germanium to change its equilibrium lattice constant but keep the harmonic force constants unaffected. The study of the relaxed interface formed from two semi-infinite bulk materials shows that lattice mismatch increases the lattice disorder in the interfacial region, as shown in Fig. 16a and decreases the adhesion energy. An abrupt change of acoustic impedance at the interface and facilitates phonon transmission into the device region, which are standing waves with zero group velocity and have little ability to travel through the device region. For the superlattice with larger period thickness, more layers are at the zone edge, leading to more oscillations in the phonon transmission curve.

Liu and Yang investigated the phonon transmission across Si/Ge superlattice sandwiched between two Si contacts. In the study, both the period thickness \( L \) (with \( L/2 \) of Si and \( L/2 \) of Ge) and the number of periods \( N \) could affect the phonon transmission across the multi-layered superlattice-like structures. With the increasing number of periods, the phonon transmission generally decreases and eventually converges. As the number of periods increase from 1 to 4, the phonon transmission is found to decrease dramatically. The reduction could be understood by the multiple interface effects, since more interfaces are more likely to scatter phonons. When the number of periods is larger than 4 periods, the phonon transmission reduction with \( N \) becomes slight and the phonon transmission converges for \( N \geq 10 \). This is attributed to the fact that superlattice phonon band has formed for the multilayer structure with 10 periods. When the number of periods are fixed at 10, more oscillations in the frequency dependent phonon transmission curve are found in the superlattice with smaller period thickness. The oscillations are originated from the band-folding-induced phonon modes at the zone edge, which are standing waves with zero group velocity and have little ability to travel through the device region. For the superlattice with larger period thickness, more modes are at the zone edge, leading to more oscillations in the phonon transmission curve.

Apart from introducing interfacial roughness to reduce the phonon transmittance, Qiu et al. also investigated the roles of aperiodicity of the superlattices. For the superlattices with aperiodicity, the transmittance monotonically decreases as more layers are added into the superlattices for phonons with frequency larger than 60 cm\(^{-1}\). The origin of the phonon transmittance reduction is similar to the that occurs in superlattice with rough interface, as the coherence is broken and phonons. If the interface roughness is included in the aperiodic superlattice, the phonon transmittance is enhanced compared with that with smooth interfaces. This can be understood as the interface mixing scatters phonons into all directions, reducing the effectiveness of phase cancellation and localization in purely one-dimensional structures.

5.3 Mode-decomposed non-equilibrium Green's function

Although the NEGF approach has been successfully applied to calculate the frequency dependent phonon transmission for many systems under the harmonic approximation, it is highly desirable to obtain the mode-specific phonon transmission. The detailed information of the mode-specific phonon transmission could be useful as inputs in
Boltzmann transport equation based method to model the phonon-interface scatterings. We will briefly discuss the implementations on calculating mode-specific phonon transmission in the NEGF calculations.

The basic idea of computing the mode-decomposed phonon transmission is to find out the vibrational patterns corresponding to specific frequency and wave vector, and then to define the Bloch matrices that generate the phase change when the phonon propagates. Based on the Bloch matrices, the mode-specific transmission could be evaluated individually.

Following Ref. 246, similar to the decimation algorithm to calculate the surface Green's function as discussed in Sec. 5.1, the contacts are divided into sublayers. The equation of motion for the $i$-th slice in the left contact is written as

$$
-H^L_{i0}\Phi_{i\lambda}^L+(\omega^L-H_{00}^L)\Phi_{i\lambda}^L-H_{0\lambda}^L\Phi_{i\lambda}^L=0.
$$

(111)

This equation is solved from the corresponding eigenvalue equation

$$
\omega^L-H_{0\lambda}^L-\lambda H_{1\lambda}^L-(H_{0\lambda}^L)^\dagger=0,
$$

(112)

where $\lambda=\exp(ikx_0)$ is the Bloch factor. Since the sublayers at the contact have translational invariance, any solution (eigenvector) of Eq. (111) satisfies $\Phi^{A_0}_{i\lambda}=\lambda^{A_0}\Phi^{L}_{i\lambda}$. Unlike the usual practice to obtain the frequency $\omega$ by solving Eq. (112), here $\omega$ is set as a constant number, and $\lambda$ is solved as an unknown variable. Then, Eq. (112) is transformed to an equation for $\lambda$

$$
-H_{0\lambda}^L u+\lambda (\omega^L-H_{0\lambda}^L) u-\lambda^2\hat{H}_{0\lambda}^L u=0,
$$

(113)

where $u$ is the right eigenvector. By solving this equation, we can distinguish every phonon mode with frequency $\omega$ through the phonon eigenvector. Since Eq. (113) is quadratic in $\lambda$, two solutions could be found. Meanwhile, for each $\lambda$, $N_i$ eigenvectors could be solved, which correspond to $N_i$ phonon modes. These $2N_i$ modes include $N_i$ right-going modes that are either evanescent waves that are decaying to the right or waves of constant amplitude that are propagating to the right, and the other $N_i$ left-going modes that can also be classified to evanescent and propagating waves. The right-going modes are labeled as $u^{m+}_{n\pm}(\lambda)$ for $n=1,\ldots,N_i$, while the left-going modes are denoted as $u^{m-}_{n\pm}(\lambda)$. Whether the modes are propagating or evanescent can be distinguished by examining the value of $|\lambda|$, which is detailed in Ref. 246. Under the time-reversal symmetric transformation, another set of modes can also be defined as $u^{m\mp}(\lambda)$.

Any wave function in the contacts can be expressed as a linear combination of the bulk modes in the contacts. It is possible to construct the so-called Bloch matrices $F$ using the solved wave functions through

$$
F^{+}_{\alpha}(\pm) U^{\pm}_{\lambda}(\pm) = U^{\pm}_{\lambda}(\pm)\Lambda^{\pm}_{\alpha}(\pm),
$$

(114)

where $\alpha=ret$ or $adv$, $U^{\pm}_{\lambda}(\pm)$ is a matrix where the columns consist of the normalized eigenvectors $u^{m+}_{n\pm}(\lambda)$ and $\Lambda^{\pm}_{\alpha}(\pm)$ is a diagonal matrix whose diagonal elements are $\lambda^{\pm}_{\alpha}(\pm)$. The Bloch matrix can be regarded as an operator that generates the phase change when phonon travels from one layer to the right layer. Instead of obtaining the Bloch matrices through Eq. (114), Ong and Zhang applied the mode-decomposed NEGF method to the study of phonon transmission across the graphene-hexagonal boron nitride interface. Fig. 17 shows the phonon transmission probability at $\omega=200$ cm$^{-1}$ on the $(k_x,k_y)$ plane. Their calculations show that the transmission probability depends strongly on the polarization and angle of incidence of the phonon mode. Yang et al. also employed this method to investigate mode-specific phonon transmission crystalline-amorphous interfaces. The amorphous layers could effectively reflect modes of frequency greater than around 3 THz, but transmit phonons below this frequency.

![Fig. 17](Image)

**Fig. 17** Transmission probability for phonon modes on the $(k_x,k_y)$ plane at $\omega = 200$ cm$^{-1}$. The transmission probability is indicated by the color bar. Reproduced from Ref. 246, with permission from American Physical Society.

### 6. Numerical solution of phonon Boltzmann transport equation

Phonon BTE plays an important role in in-depth understanding of micro- and nano-scale heat transfer. When the characteristic length scale of the system is comparable to the phonon mean-free-path but larger than its wavelength, and the phase coherence effects are unimportant, particle-based approaches based on the phonon BTE are efficient and widely used tools. In general, there are three kinds of solutions for phonon BTE: (a) Monte Carlo (MC) method; (b) lattice Boltzmann method (LBM); (c) deterministic methods. MC simulation is especially flexible for use with complex geometric configurations and can readily include different scattering mechanisms. However, it requires relative large computational costs.
to reduce the random error. As a faster solver, LBM is also easy to deal with complex structures, but the results are severely limited by the finite angular discretization in the grid model. Considering the similarity between the phonon BTE and the radiative transfer equation (RTE), deterministic solutions based on a variety of discretization techniques have been developed. Deterministic solution is faster than MC but it is more difficult to handle very complex geometries. Due to the major limitation of angular discretization of LBM, it is not widely used for phonon BTE, and thus it will not be discussed here. The MC method and deterministic method based on discrete ordinate method (DOM) will be reviewed in details in this section.

6.1 Algorithm of MC method
MC simulation has been proven to be a favorable stochastic method to solve the phonon BTE. It can well handle the transport problems involving complicated geometries, multiple scattering events, and even the heat wave effect. Two typical kinds of MC methods are often used to simulate phonon transport in nanostructures: the ensemble MC and the phonon tracing MC. The ensemble MC method simulates the trajectories of all phonons simultaneously at each time step. Early in the 1990s, Peterson used it to simulate phonon heat conduction process under Debye approximation. Afterwards, this method is used to calculate the effective thermal conductivity of various nanostructures, including composites, nanoporous silicon, etc. As for the phonon tracing MC method, the trajectories of individual phonons are simulated independently, gaining a significant reduction of computational expense.

Klitsner et al. used the phonon tracing MC simulation to study ballistic heat conduction process in the silicon crystals at an extremely-low temperature, where the influence of internal phonon scatterings can be neglected. Then, this method has been extended to simulate phonon transport processes involving internal phonon scatterings. Moreover, phonon MC simulations can also be used to investigate transient heat conduction, e.g. thermal conduction in materials during ultrashort pulse laser heating process in TDTR measurements, and heat wave phenomena, where partial derivations of distribution function to time are taken into consideration. Jean-Philippe M. Peraud et al. developed the deviational MC simulation method and studied the evolution of surface temperature in TDTR measurements. Using the similar simulation method, i.e. the phonon tracing MC simulation method, Tang et al. investigated the thermal wave phenomena in transient phonon ballistic-diffusive regime. Actually, ensemble MC simulation method is a transient method, which realizes a steady state process by a long-time transient process. For example, Lacroix et al. simulated the heat conduction in silicon nanofilms at both steady and transient state using ensemble MC simulations. In detailed manipulations, Debye approximation assuming a linear phonon dispersion, and the gray approximation which assumes phonon properties are frequency-independent, are often adopted in MC simulations. Indeed, the basic disciplines for the gray and non-gray MC simulations are exactly the same; therefore, the relevant details of phonon dispersion properties in MC simulations are not emphasized in the followings.

In this section, the basic discipline and algorithms of MC simulations for phonon heat conduction process, will be presented in detail.

Phonon tracing MC method
In practice, what we simulate is the prescribed phonon bundles but not actual phonons. The intensity of each phonon bundle is defined as \( W = E/N \), where \( E \) is the emission phonon energy per area unit time from the boundary, and \( N \) is the number of phonon bundles that we trace in MC simulations. The emitting phonon energy, \( E \), is dependent on the boundary temperature, \( T \),

\[
E = \sum_{p} \int_{0}^{N_p} v_p(\omega) C_{p} d\omega \frac{T}{4}
\]

in which \( C_{p} = \hbar \omega f_{abs}/\omega T D(\omega) \) and \( D(\omega) \) is the phonon density of states, and \( T \) refers to the boundary temperature. \( N \) must be large enough to preserve the simulation accuracy, and energy conservation is guaranteed by conserving the total number of phonon bundles.

The position vector of phonon bundle is defined as \( \mathbf{r} = [x,y,z] \), and the directional vector is \( \mathbf{s} = [\cos(\theta), \sin(\theta) \cos(\phi), \sin(\theta) \sin(\phi)] \), where \( \theta \) is the polar angle and \( \phi \) is the azimuthal angle. Besides, for non-gray simulations, phonon properties, including dispersion relations and relaxation time, which can be obtained from the first-principles method and some empirical models, should be input for initialization. With gray approximation, phonons travel with one average group velocity and the scattering rate is characterized by an average phonon MFP. In addition, heat capacity is generally approximated to be constant under conditions of small temperature variations, which can linearize the relation between phonon intensity and temperature since accurate derivation will give a result that phonon intensity is proportion to fourth power of temperature.

Here, the two most common boundary conditions, i.e., isothermal and adiabatic boundaries, are discussed. An isothermal boundary holds two functions: First, it emits phonons into the computational domain; second, it also serves as an absorbing boundary to ensure energy conservation. When a phonon bundle emits from an isothermal boundary, its intensity is given in terms of the boundary temperature, and the boundary phonon emission distributions, including angular and spatial distributions, should be obtained according to the properties of boundary. In principle, an isothermal boundary is usually set as phonon black-body in analogy to the black-body wall in photon transport, that is, phonons arriving at it will be completely absorbed. In contrast, for adiabatic boundary, all phonons that strike it will be reflected back into the computational domain. A specular parameter, \( P \), is introduced to describe the possibility of phonon specular scattering at such boundaries. It can be expressed as \( P = \exp(-\Delta T/\lambda) \), in which \( \Delta T \) is the root-mean-square value of the roughness fluctuations and \( \lambda \) is phonon wavelength. When \( P \) is equal to 1, the phonon scattering is completely specular, and we then have

\[
\mathbf{s}_r = \mathbf{s} + 2\epsilon \mathbf{s} \times \mathbf{n}
\]

in which \( \mathbf{s}_r \) is the incident direction vector, \( \mathbf{s} \) is the reflect direction vector, and \( \mathbf{n} \) is the unit surface normal vector. While \( P = 0 \) corresponds to the diffusive scattering, the reflecting direction vector should be regenerated.

Basic principle to solve the phonon Boltzmann transport equation and phonon tracing process are the same for steady state and transient cases, while in transient simulations, time discretization is required. Detailed procedures for transient phonon-tracing MC simulation are shown as Fig. 18 and following statements,

1. **Initialization**: Input phonon properties (e.g. frequency, wave vector, polarization, group velocity and relaxation time), and set total number of phonon bundles \( N \) and time step length \( dt \).

2. **Phonon bundle emission**: Draw the initial properties of a phonon bundle according to the nature of the emitting boundary.
These properties, including position $\mathbf{r}_0 = [x_0, y_0, z_0]$, traveling direction $\mathbf{s}$, polarization $p$, angular frequency $\omega$, etc., are determined by random number sampling. If the number is larger than setting number, then tracing process stops; otherwise, go to step (3).

(3) **Phonon bundle moving**: Generate dimensional phonon free path $l$, based on phonon mean free path function $\ell = -K \ln (1 - R)$, the minimum dimensional distance between phonon at current state and boundary $l_\sigma$, and dimensional phonon moving distance $l_\sigma$ at current time step. Determine the next phonon moving distance $l$ and moving time $t/l_\nu$ by choosing the minimum value among the above three distances.

(4) **Phonon scatterings and reemissions**: Determine phonon scattering events by actual phonon moving distance. If $l = l_\sigma$, then phonon-boundary scattering occurs, phonon tracing process stops for isothermal boundary and phonon is absorbed and reemitted at the adiabatic boundary (back to step (2) and continue); if $l = l_\nu$, phonon experience internal scatterings; if $l = l_\nu$, phonon moves in the previous state without being scattered.

(5) **Phonon bundle tracing termination**: Calculate the total time at the current step, if setting total time $t'$ is less than the current total time, tracing process of this phonon bundle stops; otherwise, move to step (5) and continue.

To obtained final temperature and heat flux density data with respect to time and location, location and time of phonons and phonon scatterings should all be recorded. For temperature, it can be calculated or defined based on density of scatterings and density of phonons, respectively. Heat flux is calculated based on its definition, i.e. energy (phonon bundles) transmitted across unit area during a time step.

**Ensemble phonon MC method**

Ensemble phonon MC method, which is also called directional simulation MC method, can solve the phonon BTE in kinds of forms.

![Fig. 18](image)

Fig. 18 Phonon tracing algorithm schematic for transient phonon tracing MC simulation.
including exact form with full scattering matrix\textsuperscript{264} and linearized form with or without relaxation time approximation.\textsuperscript{257} The first step to carry out this simulation is to initialize phonons in each computational cells based on frequency distribution, polarization probability, and relation between temperature and phonon number. And then, phonons drift with phonon group velocity and experience internal and boundary scatterings. Treatment on phonon boundary scatterings is the same as that in phonon tracing MC method, while the most important difference between these two methods is how to model the internal phonon scatterings. To simulate the physical process of phonon Boltzmann transport equation of exact form with full scattering matrix, and linearized form with temperature dependent relaxation time,\textsuperscript{239} phonon scatterings for next step should be determined by all phonon states, i.e. phonon distribution, at the current step. As a result, information of all phonons should be recorded at the same time, which is much different from that in phonon tracing MC simulations where phonon scatterings are independent for each phonon. It should be noted here, provided that relaxation time is temperature independent, ensemble MC simulation method is almost the same as phonon tracing MC simulation. Fig. 19 illustrates the algorithm of the ensemble phonon MC method. Basically, we divide this process to six procedures:

1) \textbf{Initialization}: Input phonon properties, discrete the simulation box into several small zones and set the total number of phonon bundles according to initial temperature distribution for each simulation zone.

2) \textbf{Phonon bundle moving}: Calculate the traveling length $\Delta r$ until the first scattering event and renew the position of phonons, $r_{\text{new}} = r + \Delta r$. In this step, all information of phonons at current state should be recorded if phonon free path or relaxation time depends on temperature and phonon distribution.

3) \textbf{Boundary scattering}: When a phonon bundle collides with a boundary at $r_b$, set $r_{\text{new}} = r_b$. If the boundary is non-absorbing, the phonon bundle should be reflected back into the domain. Then, a random number is drawn. If this random number is less than the specular parameter $P$, the boundary scattering is specular; otherwise, the boundary scattering is diffusive. If the boundary is absorbing, phonon bundle is absorbed by the boundary, and properties of this phonon bundle are then reset according to the temperature of the boundary.

4) \textbf{Phonon bundle internal scattering/reemission}: If a phonon bundle does not collide with boundaries, the phonon should experience internal scatterings at $r_{\text{new}}$. Then, we set $r_{\text{new}} = r_{\text{new}}$ and proceed to (2).

5) \textbf{Iteration process}: Since phonon free path or relaxation time depends on local temperature or local phonon distribution, iteration is needed to ensure correct temperature distribution and selection of phonon free path. Phonon and temperature distributions obtained from current simulation process are set to be the initial conditions, and continue the simulation.

6) \textbf{Termination of the simulation}: The simulation is terminated once difference between current phonon distribution and last one is smaller than setting criterion.

\section*{Issues of periodic boundary condition}

Study on the thermal conductivities of composite and phononic materials that frequently hold periodic or quasi-periodic internal nanostructures is of essential importance, due to their extensive and promising thermal related applications in engineering.\textsuperscript{265,266} For instance, researchers have demonstrated that the effective thermal conductivity of silicon films etched periodic nanoscale holes can be dramatically reduced, with only a minor effect on electrical transport ability, leading to a significant improvement of the thermoelectric performance.\textsuperscript{257} For those nanostructured materials, with their characteristic lengths comparable to the phonon mean free path, the non-Fourier effects, such as ballistic transport and boundary scattering, can lead to the geometry- and size- dependence of their effective thermal conductivities. The MC technique has been extensively used to study this issue.

When using the MC method to calculate the effective thermal conductivity, a temperature difference is usually imposed on the structure to induce a heat flux, as shown in Fig. 20, and then the effective thermal conductivity can be obtained by using Fourier's law, where $q$ is the heat flux, $L$ is the distance between these two phonon baths, and $\Delta T$ is the temperature difference. Isothermal boundary condition\textsuperscript{251} is usually used to establish the temperature...
difference, owing to its simplicity. In such case, due to the end effects, the exact effective thermal conductivity of a large-area nanostructured material can be obtained only if the simulation results will no longer vary with further increasing the number of periods between the two phonon heat baths, which results in a considerable large computational expense.

In order to reduce the computational expense, periodic boundary condition is needed in MC simulations. With using periodic boundary conditions, the effective thermal conductivity can be obtained by simulating the phonon transport process only in one or several repeating units. As shown in Fig. 20, the periodic boundary conditions can be readily imposed in the lateral direction by letting the phonons arriving at the lateral boundary reenter the domain from the opposite one, without changing momentum or energy. By contrast, in the temperature-gradient direction, the periodic boundary conditions should be set delicately to minimize the end effects.

$$N_{\text{emit}} = \frac{\pm Q + S \sum \omega h_0 \text{absorb}}{h_0 \omega_{\text{absorb}}}$$

(118)

in which $Q$ is the prescribed heat flow, whose sign depends on the boundary emission direction, $S$ is the scaling factor representing the number of phonons contained in each bundle, $S \sum \omega h_0 \text{absorb}$ is the total phonon energy leaving the computation domain, and $\omega_{\text{absorb}}$ is the average phonon frequency corresponding to the local boundary temperature, which is obtained by extrapolating the temperature profile inside the computational domain to the boundary. When a phonon bundle reaches one of the temperature-gradient directional boundaries, it is considered leaving that boundary. The pattern velocity, direction, position, and remaining flight time of phonon emission at one boundary is assumed to be the same as the pattern of phonon leaving the opposite boundary. This implementation is realized by maintaining a pool of leaving phonons, recording their velocity, direction, position, and flight time, and determining the properties of each emission phonon by randomly drawing from the pool. The pool should be refreshed after each time step. In the paper by Jeng et al., this periodic boundary condition was applied in the heat conduction simulations of nanoparticle composites. Then, following the method by Jeng et al., Hao et al. and Péraud et al. studied the thermal transport in some other periodic nanostructures.

**Periodic conditions in phonon tracing MC method**

As for the phonon tracing MC method, due to the demand of absorbing boundaries as the termination of individual phonon tracing process, the periodic boundary condition is inapplicable indeed. In order to overcome this restriction, Péraud et al. proposed that the phonon tracing process could be terminated after several scattering events with only a small effect on the simulation accuracy. This scheme is derived from their numerical observation that after several scattering events, a phonon bundle’s properties are almost completely randomized. However, it is not easy to determine the number of scattering events a phonon bundle should undergo before its tracing process can be terminated, and a rigorous proof is also lacking that this scheme can indeed approach to the exact thermal conductivity value of a nanostructure with an infinite temperature-gradient directional length where the end effects have been eliminated.

Actually, the conventional phonon tracing MC method cannot efficiently handle the transport problems in large-area nanostructured materials due to the inapplicability of periodic boundary condition. Regarding this issue, Hua et al. developed a two-step phonon tracing MC method, which greatly reduces the computation cost without degrading the accuracy. The algorithm of the two-step
phonon tracing MC method is concluded as two basic simulation steps:

(i) Phonon transport process is simulated in the initial unit to obtain the initial phonon transmittance and the phonon emission distributions at the internal virtual boundary;

(ii) According to the phonon emission distributions at the internal virtual boundary, phonon transport in the internal unit is simulated to obtain the internal phonon transmittance.

By combining the initial and the internal phonon transmittances, the total phonon transmittance and the effective thermal conductivity of the whole structure can be calculated.

MC simulation examples

This section will offer some examples using MC method to study the phonon heat conduction within various nanostructures. Particularly, the MC method was used to study the anisotropic heat conduction in the 2D periodic nanoporous films. Fig. 21a shows the structure of a 2D periodic nanoporous film that is a typical large-area nanostructured material, and thus the two-step algorithm stated above can be applied in the MC simulations. The heat flow is along the $x$-direction (along the pore axis) in the in-plane heat conduction as shown in Fig. 21b, while the heat flow is along the $y$-direction (perpendicular to the pore axis) in the in-plane heat conduction as shown.

Fig. 22 illustrates in-plane and cross-plane effective thermal conductivities of 2D periodic nanoporous silicon films with various periods ($L_p = 300, 1100$ nm) at room temperature as a function of porosity. Then, the corresponding predictive models were derived based on the simulation results,

$$\kappa_{\text{eff}} = \frac{1}{3} H \sum_{n=1}^{N} \int_0^{\infty} h \frac{\partial P_n}{\partial T} D_n(\omega) d\omega$$

with the in-plane ($\alpha_{\parallel}$) and cross-plane ($\alpha_{\perp}$) geometrical factors equal to 2.25 and 4.65, respectively.

Heat wave phenomena have been predicted for several decades after proposition of the C-V model, a well-known modification for the classical Fourier's law. By solving the phonon Boltzmann transport equation in transient state using phonon MC method, heat wave transport in phonon ballistic-diffusive regime are investigated systematically with following initial and boundary conditions,

$$\begin{align*}
q_1 &= \begin{cases} q_{\text{max}} \times \frac{1}{2}(1 - \cos(\omega t)), & t < t_0 \\
0, & t \geq t_0 \end{cases} \\
q_2 &= \begin{cases} q_{\text{max}} \times \frac{1}{2}, & t < t_0 \\
0, & t \geq t_0 \end{cases}
\end{align*}$$

where $q_{\text{max}} = 5 \times 10^7 \text{ W/m}^2$, $t_0 = 2$ ps, and $\omega_p = 3.14 \text{ rad/ps}$ shown in Fig. 23, and

$$T = T_0, \quad t = 0, 0 \leq x \leq L, \quad q = 0, \quad t \geq 0 \leq x \leq L.$$  

Temperature profiles calculated by the MC simulations and the C-V model are shown in Fig. 23. The C-V model predicts a non-dispersive dissipative thermal wave and the velocity of the wave front and the wave peak are both equal to $v_c$. However, the MC simulation predicts a dispersive dissipative thermal wave and the velocity of the wave front equates to $v_c$. Waveform of the heat pulse is not kept during the propagation process and barely influences the shapes of the temperature profiles in MC simulations as shown in Fig. 24.

6.2 Deterministic solution

The phonon BTE has been described in Eq. (3) in Sec. 2.2. To numerically solve this equation using deterministic method, several simplification are usually made. The first is the relaxation time approximation, Eq. (6). The volumetric energy density per unit frequency per unit solid angle can be obtained from distribution function by $\epsilon(\omega, r) = n(\omega, r) h_{\omega} D_{\omega}(\omega)$ (unit: $\text{J/m}^3(\text{rad/s})\text{sr}$) where $D_{\omega}$

![Fig. 22 In-plane and cross-plane effective thermal conductivities of 2D periodic nanoporous silicon films with various periods ($L_p = 300, 1100$ nm) at room temperature as a function of porosity.](image-url)
Fig. 23 Schemes of the heat pulse and the system of MC simulation for single-crystal silicon nanofilm (a) sinusoidal heat pulse; (b) rectangle heat pulse; (c) simulation system including the regimes of phonon emission and scattering.

Fig. 24 Temperature distribution profiles calculated by MC simulations and the C-V model under the stimulation of (a) sinusoidal heat pulse; (b) rectangle heat pulse. Reproduced from Ref. 270, with the permission of AIP Publishing.
denotes the phonon density of states and $p$ is the polarization. The phonon energy density $\hat{e}_n = \int e d\omega$ (unit: J/m$^2$sr), where $d\omega$ is a discrete frequency interval, is used to define the energy form of phonon BTE is then given by

$$\frac{\partial \hat{e}_{n,p}}{\partial t} + \mathbf{v}_{n,p} \cdot \nabla \hat{e}_{n,p} = \frac{\hat{e}_{n,p} - \hat{e}_{n',p}}{\tau_{n,p}} .$$

(123)

Note that here we are solving BTE to obtain the temperature and heat flux distribution over space and time, $e_{n,p}(t,\mathbf{x})$ is the phonon energy density in the position $\mathbf{x}$, in direction of propagation $\mathbf{s}$, time $t$, and per unit volume and per unit solid angle within the frequency band $d\omega$ for polarization $p$ and frequency $\omega$. The group velocity is denoted by $v_{n,p}$ and $s = v_{n,p}/\tau_{n,p}$. Because of energy conservation rule and purely randomly act of phonon for collision the integration over collision part will be zero:

$$\sum_p \int_{4\pi} e_{n,p}^0 - e_{n,p} \, d\omega d\Omega = 0$$

(124)

The equilibrium term $e_{n,p}^0$ equals to $C_v(T - T_v)$, and $T_v$ is the lattice temperature. For the gray modeling case, as a simplified version of the BTE, it equals to $\int ed\Omega/4\pi$. The boundary conditions include the isothermal, specular, diffuse, and periodic, which is similar to those described in the MC part.

BTE is a seven-dimensional nonlinear integro-differential equation, including 3 spatial coordinates, 3 wave vector coordinates, and time. The form presented above assumes isotropic wave vector spaces so the three wave vectors reduce to the direction of wave vector and the frequency of the mode. It also assumes relaxation time approximation is valid so that it can reduce to a seven-dimensional linear partial differential equation, as shown above. Deterministic solution of the BTE requires both angular and spatial discretizations. In addition, if transient form is considered, the time discretization is also needed. As the similarity between the phonon BTE and the RTE has long been recognized, deterministic solution of the BTE can therefore adopt the methodology of deterministic solution of RTE. The discrete ordinates method (DOM) has been used extensively for the numerical solution of RTE, so it has been borrowed to the phonon transport area. The DOM is also called $s_i$ approximation. The basic idea is to assume that the intensity or energy density in a solid angle does not change with direction so the angular space can be divided. As such, the entire $4\pi$ solid angle is divided into different control angles, and each has a preferred direction $\mathbf{s}_i$ as shown in Fig. 25. Using DOM, there will be one partial differential equation for each direction:

$$\frac{\partial \hat{e}_{n,i}}{\partial t} + |\mathbf{v}_{n,i}| \nabla (e_{n,i} \mathbf{s}_i) = \frac{\hat{e}_{n,i}^0 - \hat{e}_{n,i}}{\tau_{n,i}} .$$

(125)

Here $i$ denotes the energy density in the direction $\mathbf{s}_i$, and these equations are coupled through the source term $e_{n,i}^0$ by Eq. (124). The standard DOM use uniform division, as shown in Fig. 25b.

There is a shortcoming arising from the discrete angles, i.e., the ray effect. Ray effect arises from the approximation of a continuously varying angular nature of radiation by a specified set of discrete angular directions. For phonon BTE, such an unphysical effect is more pronounced in the ballistic regime (larger $Kn$) where the phonons experience less scattering. As such, when we use standard DOM, we need to use refined angular division and compute for more directions to get accurate result. This will cost much more computation time. To reduce the ray effect, we can use the control angle discrete ordinates method (CADOM). On the basis of the standard DOM, after dividing the angular space uniformly as shown in Fig. 25b, CADOM performs analytical integration over each control angle. To further reduce inaccuracy, higher order numerical integration using quadrature instead of uniform divisions can be performed. As shown in Fig. 26, although there are still some ray effect for $Kn = 10$, it is much improved as compared to DOM (see Ref. 276 for comparison). The details of various DOM approaches are presented in standard radiative heat transfer textbooks, such as the one by Modest. Apart from angular discretization, deterministic solution of the phonon BTE also require spatial discretization to further reduce the partial differential equation Eq. (125.) in to

---

**Fig. 25** In the solution of phonon BTE, the angular dependence of energy flux must be considered. Therefore, the angular domain needs to be discretized. (a) Spherical coordinates; (b) uniform discretization of solid angles.

© Engineered Science Publisher LLC 2018
algebraic equations. The most widely adopted spatial discretization is the finite volume method (FVM), due to the conservation property. Alternatively, the finite element method and meshfree method are also been used.

Recently, a finite-volume discrete unified gas kinetic scheme (DUGKS) for molecule flows ranging from continuum to rarefied regimes has been developed. Guo and co-workers used it in phonon transport. DUGKS improved the stability of the FVM by making a switch from semi-implicit scheme to explicit scheme by mathematical simplifications. Different from standard FVM which uses upwind scheme to deal with convective term, DUGKS integrates the equation along characteristic line to calculate convective term. Compared with standard FVM, DUGKS has better accuracy and stability and it has asymptotic preserving property, which is important when we calculate the region from diffusive to ballistic. The DUGKS approach, however, suffers from the limitation of Courant–Friedrichs–Lewy (CFL) condition. An approach based on the semi-Lagrange method is also developed by Zahiri et al. to solve non-gray phonon BTE, which can further overcome the CFL condition that limits DUGKS and be unconditionally stable.

Due to the complexity of BTE, the computational cost of solving BTE is quite high. Most of the researches are still focusing on the solution of 2D BTE. Many are just “toy problems”, such as the one shown in Fig. 26. Therefore, efficiency of parallel algorithm is very important to achieve device level simulation using full non-gray BTE. Different parallel strategies (domain decomposition, band decomposition, angular decomposition) have been discussed. In addition, since phonon has a broad mean free path distribution spectrum, it imposes additional difficulty to solve non-gray BTE throughout the ballistic to diffusive regime. The major reason is that the phonon intensity is strongly directional dependent in the ballistic regime and more isotropic in the diffusive regime. A few hybrid solutions have been proposed, for example, to couple BTE method with the solution of heat diffusion equation, and to couple DOM approach with spherical harmonic method (P approximation). These techniques allow for better computational efficiency. Only very recently, the non-gray solution of 3D heterostructure using an unstructured mesh has been demonstrated. The structure was discretized using 604,054 tetrahedral control volumes, 400 angles, and 40 spectral intervals (or bands), resulting in 9.7 billion unknowns. These computations were performed using 400 processors in parallel, and required 156 GB of RAM and 1.1 h per time step.

7. Hybrid/Coupling methods
The trend towards increased circuit integration on smaller electronic devices has placed a greater demand on the effective thermal management to maintain device temperatures at acceptable levels. In practical situations, a large number of micro- and nano-scale components are packaged, making the heat conduction process be a multi-scale problem. For example, the length scale of GaN high electron mobility transistors (HEMTs) ranges from ~10 nm (thickness of the barrier) to ~1 mm (die length). Accurate thermal analysis of such multi-scale heat conduction is challenging and time-consuming. The traditional Fourier's law has been widely used for chip-package-level thermal simulations, but it cannot capture nanometer-scale thermal effects. The aforementioned computational methods, such as
MD and phonon BTE, describe non-Fourier heat conduction at micro- and nano-scale well. However, their usage has been severely limited by the high computational complexity. As a result, the routine analysis of the multi-scale heat transport process by the present methods is nearly impossible. As shown in Fig. 27, first-principles calculation is an atomic level method which requires to solve Schrodinger's equation. The applicable length scale of MD simulation is limited to tens of nanometers, since it focuses on the movement of every atom or molecular. Approaches based on numerically solving phonon BTE such as phonon MC and DOM are flexible for use with larger scale structures, but they are still very difficult to use when the length scale approaches millimeters. In order to simulate the multi-scale heat conduction ranging from nanoscale to macroscopic scale, the hybrid method that couples methods suitable for different length scale has received a great deal of recent attention, its prospects for high accuracy and efficiency make hybrid method a promising development direction.

Although several studies have been published for the hybrid methods of heat conduction during the last decade, compared with these individual computational methods, research on hybrid methods is still young and needs future work. In general, there are two ways to achieve the hybrid method for heat conduction: (a) de-coupled scheme, which refers that the information is transferred by some parameters and different methods for different length scale are conducted individually; (b) coupled scheme, where the solutions of different methods will be coupled during their solving process, and the final full solution correspond to the converged results of different methods. A brief introduction to these two types of hybrid methods will be given in this section, including their basic ideas, advantages and limitations.

**De-coupled scheme**

The de-coupled scheme is mainly applicable for the coupling of microscopic and mesoscopic methods, since both the methods deal with the detailed information of heat carriers and it is easy to achieve information exchange just by several parameters. The most well-known technique is coupling first-principles calculation with MD or phonon BTE to model heat transport in solids, where the atomic potential function or force constant are the information transfer parameters, as we mentioned in Sec. 3. From this point of view, first-principles phonon BTE method can be regarded as one type of the de-couple scheme.

The results of MD simulations are very sensitive to the form of the potential function used, while those of phonon BTE heavily depend on the phonon dispersion and scattering. By coupling with first-principles calculation, the concerns about the inaccuracy of the empirical settings for these important parameters fall away. However, MD simulation is very difficult to be used for system at micrometers. Even for the mesoscopic phonon BTE, the solving procedures will be very inefficient when the length scale is larger than tens of micrometers. This limitation provides a motivation for coupling the microscopic or mesoscopic methods with macroscopic methods, which necessitates the development of the coupled scheme.

**Coupled scheme**

Instead of using some parameters to transfer information, the coupled scheme runs different methods simultaneously and the information exchange is achieved during every iteration step. For example, solving BTE for all phonons in ballistic-diffusive heat conduction is a huge task in multi-scale processes, since it is very inefficient for phonons that are in diffusive regime. It inspired people to develop hybrid phonon BTE - diffusion heat conduction equation method, which is expected to be a promising solution for multi-scale heat conduction.

Efforts have been devoted to improve the solving speed of phonon BTE by classifying the phonons and using different solvers for different kinds of phonons. Mittal and Mazumder split the phonon intensity into the ballistic and diffusive components, of which the former was solved by DOM or CADOM, and the latter was determined by invoking the first-order spherical harmonics approximation. Loy et al. established a phonon group cutoff Knudsen number (Kn = 0.1), phonon bands with low and high Knudsen numbers are solved using a modified Fourier equation and phonon BTE, separately. The two solvers are coupled by the lattice temperature. Allu and Mazumder adopted the same idea of a cutoff Knudsen number. Full angular discretization was used for the high Knudsen number bands, while for others, the spherical harmonics approximation was utilized.

These hybrid solutions of phonon BTE take advantage of the fast speed of macroscopic solver and the high accuracy of microscopic solver, making them work well for the multi-scale heat conduction process. Without much compromising on accuracy, the computation speed can be remarkably improved by several times. However, their results are strongly dependent on the phonon classification, whose criterion is still totally empirical now. In addition, phonon's properties will change by a large number of scattering mechanisms, resulting in a tedious calculation process as...
phonons have to be reclassified after every scattering.

Fig. 28 The schematic diagram of the domain dividing hybrid method.

Another way to realize the coupled hybrid method is dividing the computation domain and using different solvers for different subdomains. As illustrated in Fig. 28, the whole domain can be divided into three zones: microscopic zone (subscript 'M') where methods of micro and nano-scale are adopted and macroscopic zone (subscript 'D') where Fourier's law or simple diffusion approximation are adopted, between which is the overlap zone at which solutions are exchanged to obtain the entire full solution. For the zones covering the micro- and nano-scale components, detailed computational methods are used to characterize these non-Fourier phenomena. For other zones that are little affected by the generator of non-Fourier heat conduction, simplified treatments such as Fourier heat diffusion equation can be used to substitute the complex microscopic analysis.

In fact, such zone dividing methods have been widely adopted in the fast simulations of fluid flow and radiation. In 1995, O'Connell and Thompson firstly presented a hybrid MD-continuum simulation method. The coupling is achieved by constraining the dynamics of fluid molecules in the vicinity of the MD-continuum interface. Hash and Hassan investigated different interface conditions for the techniques that coupled direct simulation Monte Carlo (DSMC) and Navier-Stokes equation and concluded that the Marshak condition is the best choice. Wang and Jacques adopted photon MC for zones near to the photon sources or the boundary. After the photon packet enters into the ambient media, it is converted to a virtual isotropic photon source and the diffusion theory is applied. Hayashi et al. also used diffusion approximation to substitute photon MC in the high-scattering regions. The heat flux at the zone interfaces worked as an exchanger of the information. More examples of the hybrid methods for fluid flow can be found in Refs. 301-304, including the coupling between MD and macroscopic solutions.

Since phonons have many similarities with molecules and photons, the domain dividing method is supposed to perform well for heat conduction. Recently, Li et al. developed a hybrid MC-diffusion method for the ballistic-diffusive heat conduction in nano-and micro-structures. They considered that the phonon-boundary interactions are the major contributor of non-Fourier heat conduction and the middle zone of the system still follows Fourier's law. By using an alternating method and setting virtual phonon bath or specular reflection as the boundary condition for the MC zones, the results of the phonon MC and diffusion equation converged at the overlap zone. More importantly, the computation time could have a 90% reduction at the most compared with the standard phonon MC method. The domain dividing hybrid method avoids the trouble of continually classifying phonons, but owing to the challenging information exchange between the microscopic and macroscopic variables, few works are reported. In addition, the partition of the whole system will be much more complex when considering the phonon mean-free-path spectrum.

The motif of the hybrid computational method is to combine both the accuracy and efficiency to achieve the fast simulation of multi-scale heat conduction. There are two ways to achieve the hybrid methods: the de-coupled scheme where different methods are adopted in order, and the coupled scheme which refers that different solvers run simultaneously. Thanks to the increased frequency of information exchange between different methods, the coupled scheme shows greater potential for the thermal simulation of actual used devices than the de-coupled scheme. The key of the hybrid method is how to establish an information bridge between different methods in a physically meaningful way. Approaches based on classifying phonons and dividing computational domain both have gotten some achievements in the fast solution of phonon BTE, but they still have many natural limitations. More work is needed to investigate better coupling manners, so that the hybrid method is able to provide unprecedentedly convenient guidance for engineering design.

8. Conclusion

First-principles BTE is a very powerful method that can predict the thermal transport properties from the only input of atomic structures. It takes the interatomic force constants as inputs and can predict the lattice thermal conductivity of crystalline solids. This method has been applied to various materials, including numerous bulk material and low-dimensional materials. It has been recently extended to metallic materials by including electron-phonon scattering effect and electron transport. This method is relatively new and still under fast development. There are several limitations as well. First, it is still limited to crystalline materials. How to consider amorphous material, alloys and nanostructures without introducing additional fitting parameters remains a fundamental challenge. From the computational point of view, it is still quite expensive. How to extract force constants efficiently and accurately is another major limitation. Also, the prediction accuracy for the low-dimensional materials needs to be further clarified.

MD simulations possess diverse and great advantages in the modelling of thermal transport properties. EMD simulation method is powerful in the calculation of bulk thermal conductivity of various materials, and NEMD method is widely used in the calculation of thermal conductivity at finite size and its length dependence, as well as the interfacial thermal resistance. The underlying mechanisms of thermal transport can also be explored from MD simulations. For example, vibration frequencies from density of states, phonon localization from participating ratio, phonon relaxation time and mean free path from SED simulations, and single-mode phonon transport from wave packet method. However, MD simulation for modeling thermal transport still has some limitations to be further developed. For instance, the modeling of the quantum effect at low temperature, lacking of empirical potential for complex materials, and insufficient modeling capability at large scale for real problems and applications, to name a few. Those limitations should be studied and resolved in the further developments of MD simulations. The analysis methods also have limitations in the applications, such as...
the SED simulations limited to the crystalline, and wave packet method only valid at the low temperature. The analysis of thermal transport in amorphous from mode/frequency perspective still a problem that deserve to be resolved, such as in polymers.

Non-equilibrium Green’s function is a useful tool to explore the phonon transport in nanostructures under the quasi-ballistic regime. It has been successfully applied to low-dimensional materials, interfaces between dissimilar materials, multilayer structures, and the dependence of the detailed atomic configurations on the phonon transmission and thermal conductance across these structures has been investigated. The implementation of calculating mode-decomposed phonon transmission makes it possible to integrate NEGF with BTE-based methods to model phonon transport in more complicated nanostructures.

Phonon BTE can be solved by MC method or DOM method. MC technique is a stochastic method to solve the phonon BTE with considerably simple algorithms, and it has been widely used to study nanoscale/ultrafast heat conduction within various nanostructures, including films, wires, nano-porous and composite structures, etc., at mesoscopic level, in both transient and steady states. DOM method is a deterministic in comparison with MC. It is computationally faster than MC but it is more difficult to handle complex geometry. To numerically solve BTE, microscopic information including phonon properties and phonon interface transmittance are required, which can be obtained by MD simulations or first-principles calculation.

Hybrid/coupling methods become highly desired since there is a tremendous increase in computational complexity and cost when the length scale of the problem spans several orders of magnitude, necessitating the development of new methods which are able to get a good balance of the accuracy and efficiency. The widely used way is to combine micro-nanoscale methods with the macroscopic ones, which generally can be achieved by two schemes: (1) the de-coupled scheme where different methods are adopted in order, and (2) the coupled scheme which refers that different solvers run simultaneously. Although the performance of hybrid methods have been verified in the last decades, it is still a young computational method which needs further developments and applications.

Conflict of interest
There are no conflicts to declare.

Acknowledgements
The contributions from Dr. Cheng Shao, Dr. Han Xie, Zhen Tong, and Yue Hu from Shanghai Jiao Tong University, Zhongwei Zhang at Tongji University, and Yuchao Hua, Jihang Zou, Daosheng Tang, Hanliang Li at Tsinghua University are gratefully appreciated. H.B. acknowledges the support from National Natural Science Foundation of China (Grant No. 51676121). J. C. acknowledges support from the National Key Research and Development Program of China (Grant No. 2017YFB0406000), the National Natural Science Foundation of China (Grant Nos. 51576153 and 11334007), Shanghai Science and Technology Committee (Grant Nos. 18JC1410900 and 17ZR1448000), and National Youth 1000 Talents Program of China. X.G. acknowledges the support from the National Natural Science Foundation of China (NSFC) (Grant No. 51706134) and the Shanghai Pujiang Program (Grant No. 17PJ1404500). B.C. is grateful to the National Natural Science Foundation of China (Grant No. 51825601, 51676108 and 51628602), Science Fund for Creative Research Groups (No. 51621062).

References