

Physics-informed neural networks for solving multiscale mode-resolved phonon Boltzmann transport equation

R. Li ^a, E. Lee ^{b, **}, T. Luo ^{a, c, d, *}

^a Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, 46556, USA

^b Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do, 17104, South Korea

^c Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN, 46556, USA

^d Center for Sustainable Energy of Notre Dame (ND Energy), University of Notre Dame, Notre Dame, IN, 46556, USA

ARTICLE INFO

Article history:

Received 22 March 2021

Received in revised form

28 April 2021

Accepted 30 April 2021

Available online 5 May 2021

Keywords:

Phonon transport

Deep neural networks

Boltzmann transport equation

Multiscale heat transfer

ABSTRACT

Boltzmann transport equation (BTE) is an ideal tool to describe the multiscale phonon transport phenomena, which are critical to applications like microelectronics cooling. Numerically solving phonon BTE is extremely computationally challenging due to the high dimensionality of such problems, especially when mode-resolved properties are considered. In this work, we demonstrate the use of physics-informed neural networks (PINNs) to efficiently solve phonon BTE for multiscale thermal transport problems with the consideration of phonon dispersion and polarization. In particular, a PINN framework is devised to predict the phonon energy distribution by minimizing the residuals of governing equations and boundary conditions, without the need for any labeled training data. Moreover, geometric parameters, such as the characteristic length scale, are included as a part of the input to PINN, which enables learning BTE solutions in a parametric setting. The effectiveness of the present scheme is demonstrated by solving a number of phonon transport problems in different spatial dimensions (from 1D to 3D). Compared to existing numerical BTE solvers, the proposed method exhibits superiority in efficiency and accuracy, showing great promises for practical applications, such as the thermal design of electronic devices.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

Effective thermal management has long been critical to the design and advancement of electronic devices. With the rapid development of micro/nano technology, the characteristic length of such devices becomes comparable or even smaller than the mean free path of phonons, which are the dominant thermal energy carriers in semiconductors. In these conditions, thermal transport is not purely diffusive, and thus it cannot be accurately described by the conventional Fourier's law [1–3]. To describe the multiscale thermal transport process from ballistic to diffusive regions, phonon Boltzmann transport equation (BTE) has been widely used when phase coherence effects are not important [4,5]. Since atomistic level simulation techniques, such as molecular dynamics

and first-principles calculations, are computationally impractical for device-level thermal analysis, mesoscopic approaches based on the solution of phonon BTE provide a balance between computational complexity and accuracy.

However, solving the phonon BTE is a challenging task because of its high dimensionality. The BTE for phonons is a nonlinear integro-differential equation with eight independent variables: phonon frequency, phonon polarization, time, three spatial coordinates (x, y, z directions), and two directional angles (polar angle and azimuthal angle). To make the solution to the phonon BTE more computationally tractable, simplifications such as assuming all phonon modes with the same properties (i.e., gray model) are commonly used [6–9]. However, such treatment can lead to significant inaccuracy in the solutions [10,11]. This is because different phonon modes can have a large span of mean free paths, and they behave differently at a given physical length scale. To address this issue, many numerical BTE solvers have been proposed, such as the Monte Carlo (MC) method [12–16], the lattice Boltzmann method [17,18], and deterministic discretization-based methods [19–22]. While the MC method has achieved great success in accounting for

* Corresponding author. Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, 46556, USA.

** Corresponding author.

E-mail addresses: eleest@khu.ac.kr (E. Lee), tluo@nd.edu (T. Luo).

dispersion, polarization and various scattering mechanisms, it suffers from stochastic statistical uncertainty and becomes prohibitively inefficient in or near the diffusive regime due to the restrictions on time step and grid size [13,14]. Improvements in MC methods have recently been made for solving 3D systems based on variation reduction techniques [15,23]. Different lattice Boltzmann methods have also been used in recent years, but there are still unphysical predictions reported for cases in the ballistic regime [24,25]. The discrete ordinate method (DOM), which solves the BTE directly using finite element or finite difference methods, discretizes the angular domain into small solid angles to capture the highly non-equilibrium phonon distribution. However, DOM and its variants usually show slow convergence near the diffusive limit and require large memory for solutions [26,27]. To solve the phonon BTE consistently over a wide length scale range, characterized by the Knudsen number (Kn , the ratio of the phonon mean free path to the characteristic length of the system), the hybrid ballistic-diffusive method [28,29] introduces a cutoff Knudsen number to separate frequency intervals for which the DOM or the Fourier's Law should be used. Although it can be applied to problems at different length scales, the proper selection of the cutoff Knudsen number remains a question and the improved accuracy comes at the cost of increased computational time. Recently, the discrete unified gas kinetics scheme (DUGKS) [30,31] and implicit kinetic scheme (IKS) [32] have been shown to work efficiently and accurately for low-dimensional thermal transport problems, but they have not been used for solving phonon BTE in 3D geometries.

It can be seen that all the aforementioned numerical schemes have some deficiencies, and few of them can be universally accurate for a wide range of length scales. Moreover, most methods have been only demonstrated for simple 1D or 2D systems, because they usually require massive computational resources for 3D geometries. For instance, while deterministic numerical methods such as IKS can solve 1D and 2D problems accurately within a few minutes [32], it can take several hours to simulate a 3D system for a single time step on a parallel machine with 400 processors [22]. Therefore, there is a pressing need to develop a numerical method that is accurate, efficient, easy to use and able to deal with high-dimensional cases.

Recently, machine learning (ML) methods have been leveraged to help predict thermal properties at the atomistic scale (1–10 nm), such as the development of ML potentials for molecular dynamics simulations [33–36]. However, these efforts are not sufficient to model thermal transport phenomena in the meso- and micro-scales (e.g., 10 nm - 100 μ m), which are more relevant to device applications. In the meantime, as deep neural networks (DNNs) possess the capability of accurately approximating any continuous function [37–39], they can be used to approximate solutions of partial differential equations (PDEs). Leveraging this property, physics-informed neural networks (PINNs) have emerged recently [40–43], which incorporate the PDE residuals into the cost function

governing PDEs are known, the solutions can be learned in a physics-constrained manner without the need for any labeled training data, which is known as a data-free ML method. PINNs have been successfully employed to simulate forward and inverse problems for a variety of PDEs [44–49]. Compared to the conventional mesh-based methods (e.g., finite element method), PINNs avoid discretizing the PDE by taking advantage of the automatic differentiation of DNNs [50]. It has also been shown that in many problems DNN as a universal function approximator can overcome the curse of dimensionality, since the number of parameters of a DNN grows at most polynomially in both the reciprocal of the prescribed accuracy ϵ and the dimension d [51,52]. In addition, PINNs have been shown capable of learning solutions of PDEs in parameterized spaces (e.g., variable initial/boundary conditions, geometry, equation parameters) [48,53]. This feature is especially valuable for optimization applications since a numerical procedure for solving PDE has to be performed each time any parameter is changed using conventional solvers – a problem also applicable to BTE solvers. In multiscale thermal transport problems, one such variable parameter could be the characteristic length. However, research efforts for solving PDEs with PINNs have been mainly in the fluid dynamics field aiming to solve Navier-Stokes equations, which usually have three input parameters. The mode-resolved phonon BTE, which can have eight variables in 3D problems, may benefit even more from the PINN scheme, but it has never been studied.

In this work, we demonstrate the use of PINNs to efficiently solve mode-resolved phonon BTE for multiscale thermal transport problems. In particular, a PINN framework is devised to predict the phonon energy distribution by minimizing the residuals of governing equations and boundary conditions, without the need for any labeled training data. Moreover, PINN can be trained to parameterize the solutions with varying geometric parameters, such as the characteristic length scale. This enables our method to handle structures over a wide range of length scales after a single training. The effectiveness of the present scheme is demonstrated by solving a number of phonon transport problems in different spatial dimensions (from 1D to 3D). Compared to conventional numerical BTE solvers, the proposed method exhibits superiority in efficiency and accuracy, showing great promises for practical applications, such as the thermal design of electronic devices.

2. Results

2.1. PINNs for stationary phonon BTE

It can be shown that for a system without internal heat source at steady state, the multiscale thermal transport problem can be described by the BTE with the single mode relaxation time approximation (see the Methods section for derivation):

$$R(e(\mathbf{x}, \mathbf{s}, k, p, \mu)) = 0 := \begin{cases} \mathbf{v} \cdot \nabla e - \frac{e^{eq} - e}{\tau} = \mathbf{v} \cdot \nabla e + \frac{e^{neq}}{\tau} = 0 \\ \nabla \cdot \mathbf{q} = \nabla \cdot \sum_p \int_0^{\omega_{max,p}} \int_{4\pi} \mathbf{v} e d\Omega d\omega = 0, \end{cases} \quad \mathbf{x}, \mathbf{s}, k, p \in \Gamma, \quad \mu \in \mathbb{R}^d, \quad (1)$$

and train the solutions using fully-connected DNNs. When the

where e represents phonon energy distribution function of

variables (spatial coordinates \mathbf{x} , directions \mathbf{s} , wave number k , and phonon polarization p) in domain Γ and parameters μ , τ is the relaxation time of phonon with k and p , and $e^{neq} = e - e^{eq}$ is the non-equilibrium part of the distribution function. Here, parameters μ include the characteristic length of the simulation system and phonon mean free path. The heat flux \mathbf{q} is calculated by integrating e over the solid angle space (Ω) and frequency space (ω , p). The solutions of e can be determined when suitable boundary conditions are prescribed,

$$\mathcal{B}_i(\mathbf{x}, \mathbf{s}, k, p, e, \mu) = 0, \quad \mathbf{x}, \mathbf{s}, k, p \in \Gamma_b, \quad \mu \in \mathbb{R}^d, \quad (2)$$

where \mathcal{B}_i are operators that define various boundary conditions, and Γ_b denotes the boundary region. In the Methods section, we show three typical boundary conditions encountered in phonon BTE, including the isothermal boundary, the diffusely reflecting boundary, and the periodic boundary.

For a given set of parameters μ , e can be obtained by conventional numerical BTE solvers, which usually involve mesh generation and iteratively solving large linear/nonlinear systems. This process can be extremely computationally costly for mode-resolved phonon BTE given the large number of input variables. In addition, solving problems with variable geometries is further challenging with these methods as any change of the geometry requires regeneration of the mesh and performing the calculations over again. Thus, we have not seen many studies adopting mode-resolved models for 3D solutions.

To enable fast predictions of steady-state multiscale thermal transport, we develop a PINN model to approximate the solutions of phonon BTE (Eq. (1)) with the length scale of simulation domain as an additional parameter included in parameters μ . It is the inclusion of such physical quantities as input variables that makes our PINN scheme capable of solving BTE in the parametric setting. This PINN framework, as shown in Fig. 1, is expected to offer a rapid online evaluation of the phonon transport with any given characteristic dimension of the target structure after offline pre-training. The input layer is composed of \mathbf{x} , \mathbf{s} , k , p , mean free path $\Lambda = |\mathbf{v}|\tau$,

and length scale L . Apart from the length scale, the mean free path is also included as an input parameter because it impacts Eq. (1) for each phonon mode through relaxation time. Given that the magnitudes of e^{eq} and e^{neq} can be very different when the frequency-dependent Knudsen number $Kn = \Lambda/L$ is extremely small (e.g., $Kn \sim O(10^{-4})$), two fully-connected DNNs are used to approximate the equilibrium and non-equilibrium distribution functions, respectively. Each sub-network maps the inputs to a target output, through several layers of neurons consisting of either affine linear transformations or scalar non-linear activation functions. The DNNs are trained by minimizing a weighted sum of residuals of Eqs. (1) and (2) as follows:

$$\mathcal{L}(\mathbf{W}, \mathbf{b}) = \left\| \mathbf{v} \cdot \nabla e - \frac{e^{eq} - e}{\tau} \right\|^2 + \|\nabla \cdot \mathbf{q}\|^2 + \sum_i \|\mathcal{B}_i\|^2, \quad (3)$$

where \mathbf{W} and \mathbf{b} refer to the weights and biases of the entire network, and $\|\cdot\|$ is L_2 norm. The boundary conditions are formulated as additional penalty terms into the loss function and imposed in a soft manner, i.e., formulated as part of the loss function instead of directly imposed in the solution. An optimal set of network parameters is identified by minimizing this composite loss function,

$$\mathbf{W}^*, \mathbf{b}^* = \operatorname{argmin}_{\mathbf{W}, \mathbf{b}} \mathcal{L}(\mathbf{W}, \mathbf{b}). \quad (4)$$

It is noted that several first derivatives of e with respect to spatial coordinates \mathbf{x} are required for constructing the PDE residuals. These derivatives can be computed accurately and efficiently by automatic differentiation in the DNN, which uses the chain rule to back-propagate derivatives from the output layer to the inputs. This approach yields a smooth representation of the solutions that can be evaluated with any input in a certain domain of interest. The above optimization problem is solved by stochastic gradient descent algorithms. The specific DNN architecture and training details are included in the Methods section.

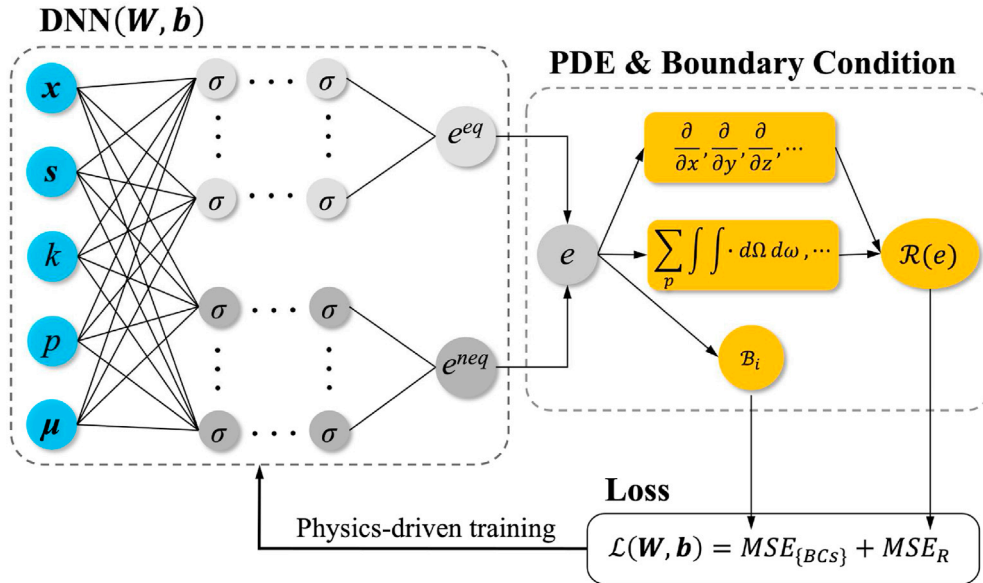


Fig. 1. Schematic of PINN framework for solving stationary mode-resolved phonon BTE. Two DNNs are employed to approximate the equilibrium (e^{eq}) and non-equilibrium (e^{neq}) parts of the phonon energy distribution, respectively. Inputs include spatial vector \mathbf{x} , directional unit vector $\mathbf{s} = (\cos\theta, \sin\theta\cos\phi, \sin\theta\sin\phi)$ (θ is the polar angle and ϕ is the azimuthal angle), wave number k and polarization p . μ represents additional parameters, which include modal mean free path Λ and system characteristic length L in this study. σ represents the activation function, which is set to be the Swish activation function in this work. The loss function of PINN contains residuals of the PDEs and boundary conditions on a sample of collocation points in the domain of interest. The hyperparameters in the DNNs are initialized by He's method [54] and learned by minimizing the total loss.

2.2. Model systems

To evaluate the performance of the proposed PINN scheme, several steady-state thermal transport problems are investigated at different characteristic length scales, including 1D cross-plane, 2D in-plane, 2D square and 3D cuboid heat conduction problems. We use single crystalline silicon as a model material system, which is the most representative semiconductor in electronics, but the proposed PINN scheme is applicable to other materials. The phonon dispersion relation of silicon in the [100] direction is used, and isotropy is assumed [55]. Only one longitudinal acoustic (LA) and two degenerate transverse acoustic (TA) phonon branches are considered because the optical branches contribute little to the thermal transport [14]. We note that including more branches is possible and only involves expanding the input space to the PINN. The phonon dispersion relation and relaxation times are determined based on Refs. [56,57] (see Supplementary Note 1 for the phonon dispersion and relaxation times considered). As we assume that the silicon simulated is not heavily doped like those in real electronics, electron-phonon interactions are not considered important in this study, but this effect can be easily included by adding its influence in the phonon relaxation times [58]. For each phonon branch, the wave vector space $k \in [0, 2\pi/a]$ is equally discretized into N_B frequency bands, where $a = 5.431$ Å is the lattice constant for silicon. The corresponding bulk thermal conductivity at 300 K for 10 frequency bands is 145.6 W/(m·K), which is close to the literature value [59], and thus we set $N_B = 10$ in all subsequent calculations. In all cases, the reference temperature T_{ref} is set as 300 K. The training and testing information about numerical experiments are summarized in Table 1. For a given length scale L , N_x is the number of interior points in the spatial domain, and N_s is the number of solid angles. N_l represents the number of length scale parameters L sampled in a certain range in Table 1. The training and validation losses of all cases are shown in Table 2, where the training loss is defined in Eq. (3) after nondimensionalization and the validation loss is the total loss evaluated in testing with the settings shown in Table 1.

2.3. 1D cross-plane phonon transport

We first consider the heat conduction across a silicon thin film (see inset in Fig. 2(a)). Phonon-mediated heat conduction in thin films plays a crucial role in nanoscale devices, and much effort has been devoted to the measurement of cross-plane thermal conductivity of thin films [60,61]. Computationally, conventional numerical methods have been proved capable of yielding good predictions of cross-plane heat conduction [30,32,62]. Given that the lateral dimensions of the film are much larger than the thickness, the heat conduction reduces to an 1D problem governed by an 1D phonon BTE with two isothermal boundary conditions (see the Methods section). The thickness of the film is L , and the left (right) boundary is fixed at $T_L = T_{ref} + \Delta T/2$ ($T_R = T_{ref} - \Delta T/2$) with ΔT

Table 2

Training and validation losses of numerical experiments.

Case	Training Loss	Validation Loss
1D cross-plane	2.0×10^{-4}	1.4×10^{-3}
2D in-plane	5.6×10^{-4}	4.5×10^{-3}
2D square	2.4×10^{-2}	1.6×10^{-2}
3D cuboid ^a	6.5×10^{-3}	8.8×10^{-3}
3D cuboid ^b	9.8×10^{-3}	4.9×10^{-3}

set to 1 K.

To train our PINN model, the spatial domain is equally discretized with 40 interior training points, and 16-point Gauss-Legendre quadrature is used for the direction $\mathbf{s}_x = \cos\theta$ (Table 1). The setting of $N_x = 40$ for the training has been validated to ensure the model accuracy (Supplementary Note 2). For the purpose of learning the parametric solutions for structures of varying characteristic dimensions, a vector of log values of thickness L is included as a part of the input (μ in Fig. 1) to the network. The model is trained on five values of L in the range between 10 nm and 100 μ m. After training, the temperature and heat flux can be evaluated at new length scales given the interpolation ability of DNNs.

Fig. 2(a) shows the dimensionless temperature profiles $T^* = (T - T_R)/(T_L - T_R)$ with different thicknesses L . The analytical solutions obtained by the method of degenerate kernels [63] are included for comparison. It is observed that the temperatures evaluated by the PINN model agree almost perfectly with analytical solutions. The discrepancy of these two models is less than 0.9%. For cases with small thicknesses, PINN successfully captures the temperature slips near the boundary because of the non-equilibrium phonon transport. We also compute the dimensionless thermal conductivities (k_{eff}/k_{bulk}) at different length scales, where $k_{eff} = qL/\Delta T$ is the effective thermal conductivity if Fourier's law

were used, and $k_{bulk} = \frac{1}{3} \sum_p \int_0^{\omega_{max,p}} C|\mathbf{v}|^2 \tau d\omega$ represents the bulk thermal conductivity in the diffusive limit. As shown in Fig. 2(b), good agreement is observed for all cases (error < 0.3%). While only trained on five parameter points in characteristic length (i.e., thickness in this case), this model is able to provide accurate predictions of thermal conductivity at unseen input thicknesses.

As for the computational cost, the training time of one PINN scheme in this case is approximately 48 min as shown in Table 1, but the online evaluation takes only a few seconds depending on the input size. It is also noted that we have only tested the model in the training range ($L \in [10^{-8}, 10^{-4}]$), and the extrapolation is usually much more challenging. However, the generalization ability is not critical to this data-free learning framework since it can be freely trained by drawing collocation points in the range of interest. In addition, the test error can be further reduced with more sample points used in training (Fig. S1).

Table 1

Training and testing information about numerical experiments. Walltimes are on a NVIDIA GeForce TITAN Xp GPU. Two models are trained for 3D geometries with different ranges of length scale L , as denoted with superscript a and b.

Case	Training				Testing				L (m)
	N_x	N_s	N_l	Walltime (h)	N_x	N_s	N_l	Walltime (s)	
1D cross-plane	40	16	5	0.80	40	64	17	2.11	$[10^{-8}, 10^{-4}]$
2D in-plane	300	144	5	4.83	1600	1024	17	24.72	$[10^{-8}, 10^{-4}]$
2D square	450	144	4	6.62	2601	576	7	9.57	$[10^{-8}, 10^{-5}]$
3D cuboid ^a	1200	144	3	15.52	132,651	576	5	344.54	$[10^{-8}, 10^{-6}]$
3D cuboid ^b	1200	64	3	8.68	132,651	576	5	351.64	$[10^{-5}, 10^{-3}]$

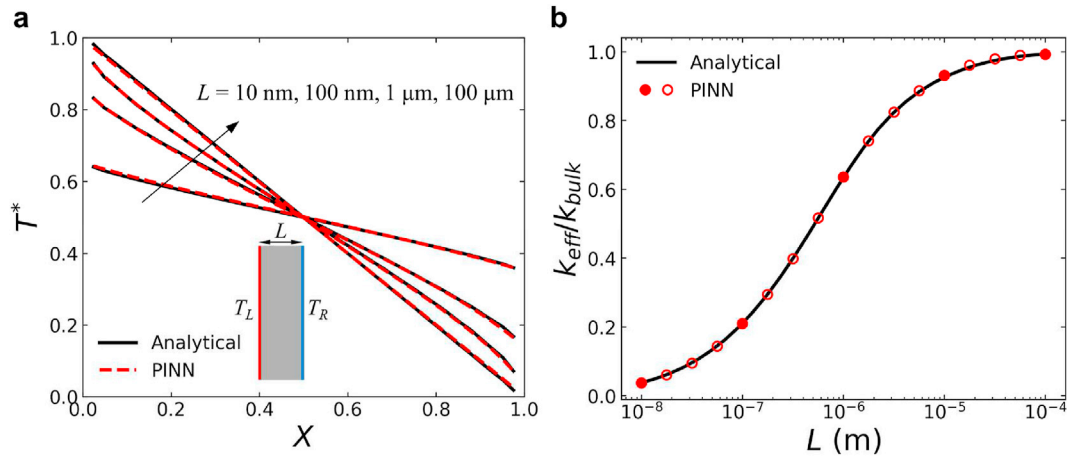


Fig. 2. Results of 1D cross-plane phonon transport. (a) Dimensionless temperature profiles of silicon thin films with different thicknesses ($L = 10 \text{ nm}, 100 \text{ nm}, 1 \mu\text{m}, 100 \mu\text{m}$), where $T^* = (T - T_R)/(T_L - T_R)$ and $X = x/L$. Inset shows the film geometry and boundary conditions. The black solid lines are analytical solutions to phonon BTE. (b) Effective thermal conductivity normalized by the bulk thermal conductivity at different film thicknesses. The filled circles represent the parameter points used for training, while the open circles are predicted points not included in training. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

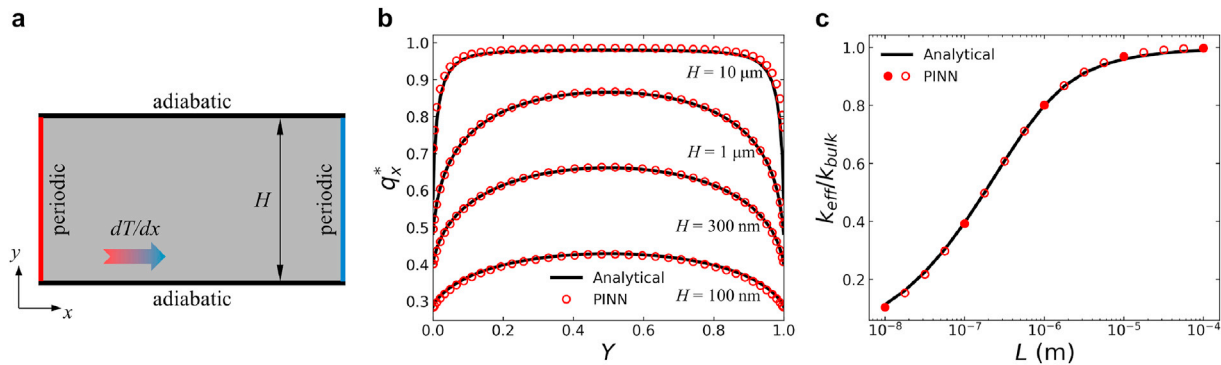


Fig. 3. Results of 2D in-plane phonon transport. (a) Schematic of 2D in-plane phonon transport case. (b) Distribution of dimensionless x-directional heat flux along the y-axis in thin films with different thicknesses, where $q_x^* = q_x(Y)/q_{bulk}$ and $Y = y/H$. From bottom to top, the PINN predictions are shown as red circles for $H = 100 \text{ nm}, 300 \text{ nm}, 1 \mu\text{m}, 10 \mu\text{m}$. The black solid lines denote analytical solutions by the Fuchs-Sondheimer theory [65]. (c) Effective thermal conductivity normalized by the bulk thermal conductivity at different length scales. The filled circles represent the parameter points used for training, while the open circles are predicted points not included in training. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.4. 2D in-plane phonon transport

The next case is in-plane thermal transport in two spatial dimensions. We consider the setup depicted in Fig. 3(a), where a silicon film is bounded by two adiabatic walls with a distance of H in the y -direction. A uniform temperature gradient is imposed along the x -direction. Diffusely reflecting boundary condition (see the Methods section) is applied for the two adiabatic walls, while the left and right boundaries are treated as periodic boundaries. In the spatial domain, 300 Sobol points for a square-shape regime are used as the interior training points, along with 30 equidistant points used for each boundary (Supplementary Note 3). To avoid the ray effect [64] induced by the conventional S_N quadrature, we discretize $\cos\theta$ in $[-1, 1]$ and the azimuthal angle ϕ in $[0, \pi]$ (due to symmetry) into 12×12 discrete angles using Gauss-Legendre quadrature. In this 2D problem, the characteristic length H is treated as an input variable parameter, and we train a PINN model to predict the phonon transport at the length scale ranging from 10 nm to 100 μm .

Fig. 3(b) shows the dimensionless x-directional heat flux $q_x^* = q_x(Y)/q_{bulk}$ for different H values, where $q_{bulk} = -k_{bulk} \cdot dT/dx$. The results are consistent with the analytical solutions by the Fuchs-

Sondheimer theory [65] over a wide range of length scales, with a prediction error less than 2.4%. We also compare the effective

thermal conductivity $k_{eff} = -(dT/dx)^{-1} \int_0^1 q_x(Y) dY$ with the benchmarks. As shown in Fig. 3(c), predictions by PINN agree well with the analytical solutions (error < 1.6%), and again this method captures the thermal conductivity variation due to the size change. It should be mentioned that all the evaluations are performed on a much denser grid mesh than the one used in training. The training takes about 4.83 h, while the evaluation costs less than 25 s for the total of 17 different H values with a 40×40 spatial mesh and a 32×32 angular mesh (see Table 1).

2.5. 2D square phonon transport

We further consider a case of a 2D square silicon as illustrated in Fig. 4(a). The temperature at the top boundary is maintained at T_h , while other boundaries are held at a lower temperature T_c . The temperature difference is set as $\Delta T = T_h - T_c = 1 \text{ K}$, and the isothermal boundary conditions are imposed on all the boundaries.

In this case, we employ a non-uniformly sampled spatial domain for improved accuracy near the top boundary (Supplementary Note

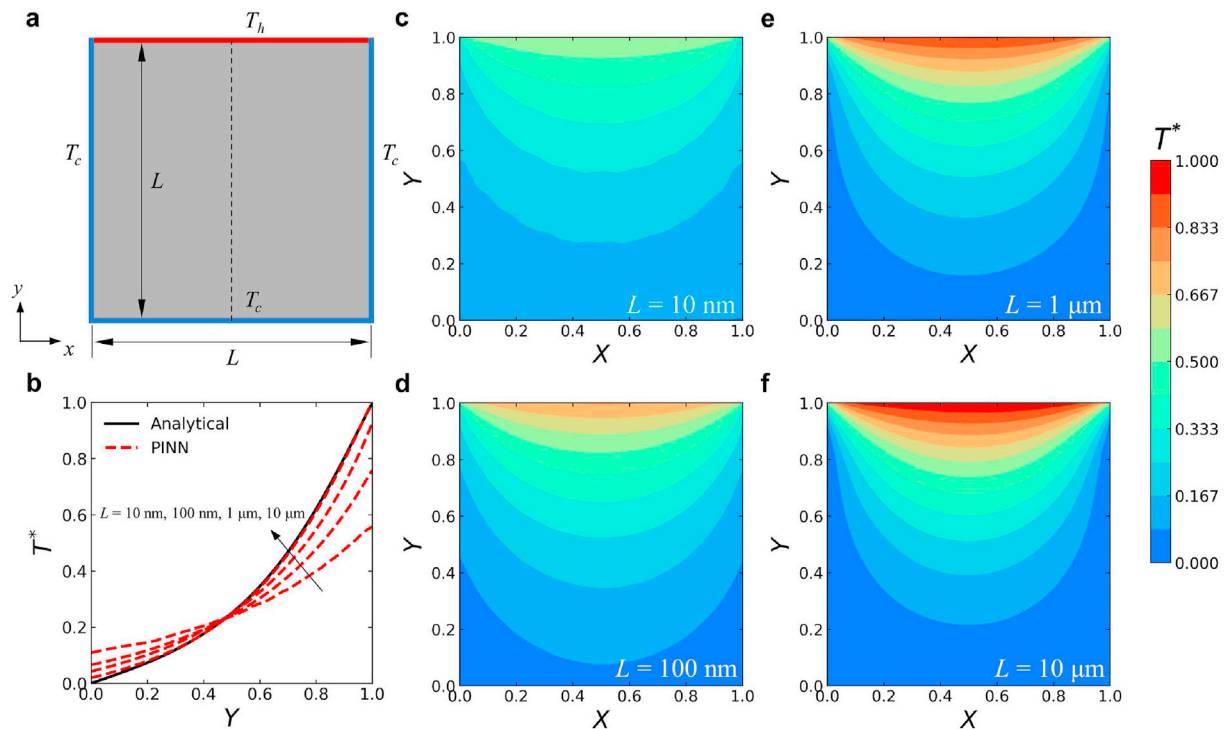


Fig. 4. Results of 2D square phonon transport. (a) Schematic of 2D square phonon transport case. (b) Non-dimensional temperature profiles along the vertical centerline (dashed line in (a)) at different length scales. Analytical solution is derived based on Fourier's law (i.e., solution in the diffusive limit). (c–f) Contours of dimensionless steady-state temperature at different length scales. X and Y are normalized spatial coordinates.

3). Similarly, the training is conducted on only four geometries with length L ranging from 10 nm to 10 μm , and the trained model can be tested at any input length scale in this range. It is noted that the validation loss is lower than the training loss (Table 2). This is because we use non-uniformly distributed interior points with local refinement near the top boundary in training, while in testing we adopt a uniform mesh.

Although there is no available numerical result for direct comparison in this case, we show the temperature profiles along the vertical centerline at $x = 0.5L$ in Fig. 4(b), together with the analytical solution to the 2D steady-state Fourier heat equation (i.e., solution in the diffusive limit). It is observed that the result for the 10 μm case is nearly identical to the analytical solution. We note that the average mean free path of phonons in silicon is around 300 nm, which is much shorter than 10 μm . Thus, the 10 μm case is supposed to be close to the diffusive limit. Fig. 4(c–f) show the 2D dimensionless temperature contour $T^* = (T - T_c)/\Delta T$ at different length scales. We see obvious temperature slips near the hot boundary for the cases with smaller L , and as L increases the slip decreases.

2.6. 3D cuboid phonon transport

To assess the potential applicability of our PINN scheme to real devices, a 3D geometry is considered as depicted in Fig. 5(a). The test geometry is a silicon block of dimensions $L \times L \times 0.5L$, with a circular hot spot of Gaussian temperature distribution on the top surface, while the other surfaces are maintained at a lower temperature T_c . Such a hot spot may correspond to a hot spot in a semiconductor layer due to the Joule heating effect. The Gaussian temperature distribution has T_{max} at the center of the top surface, and the temperature smoothly decreases to T_c ($\Delta T = T_{\text{max}} - T_c = 1 \text{ K}$). The standard deviation of the distribution is set as $L/6$, with

the full width at half maximum (FWHM) to be around $0.4L$. Two PINNs are trained for the parametric learning of the multiscale thermal transport problem. One model is for the edge length L between 10 nm and 1 μm , and the other is for L between 10 μm and 1 mm. Each model is trained on three sample geometries of varying sizes. We note that separating the training for two smaller dimension regions can reduce the total number of collocation points and improve the prediction accuracy. 12×12 discrete angles are set for the training of the smaller geometries (3D cuboid^a in Table 1), while 8×8 angles are used for the larger geometries (3D cuboid^b in Table 1). This is because the phonon energy distribution is fairly isotropic in the diffusive regime, and we can use fewer angles without deteriorating the accuracy.

The steady-state temperature contour for the structure of length $L = 1 \text{ mm}$ is shown in Fig. 5(b), which is evaluated on a computational domain with much more points compared to the one used in training (see Table 1). It is observed that PINN successfully reproduces the Gaussian temperature distribution on the top surface (Fig. 5(c)). In Fig. 5(d–i), we show temperature contours in the central planes (outlined by dashed lines in Fig. 5(a)) at different length scales. It is noted that 3D cases have rarely been investigated using BTE in previous studies, because it requires enormous computational resources. Thus, there is no solution available for comparison in this case. Nonetheless, we adopt a simple PINN model to obtain the solution of 3D heat conduction under the same boundary conditions based on the Fourier's law, as shown in Fig. 5(i). This result can be treated as the benchmark solution in the diffusive limit since the final training loss is as low as 3×10^{-4} . Comparing PINN prediction in Fig. 5(h) with the benchmark in Fig. 5(i), we find the error to be less than 1.4% (Fig. S3) as the 1 mm case should already be in the diffusive limit. Therefore, we expect high accuracies for geometries of smaller sizes as well. We note that it only takes a few minutes to solve the BTE using the PINN scheme

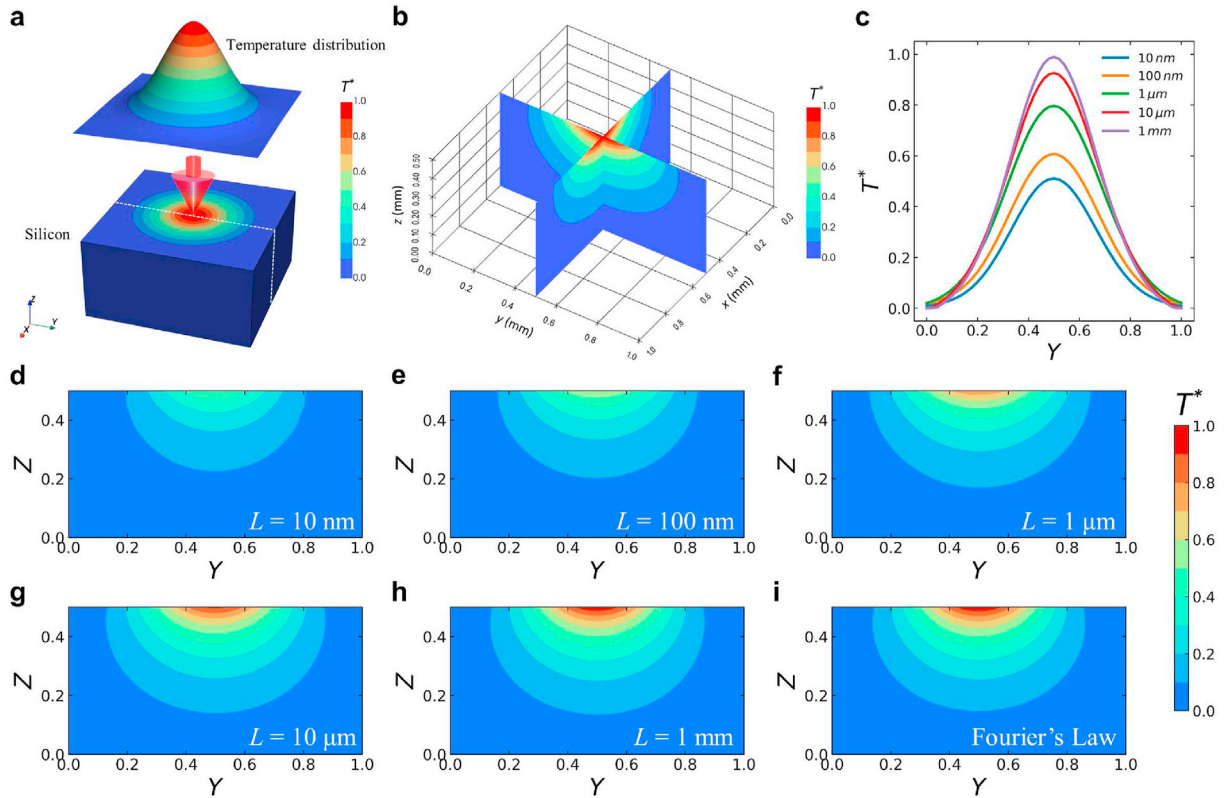


Fig. 5. Results of 3D cuboid phonon transport. (a) Schematic of the 3D phonon transport case. The system is a silicon block of dimensions $L \times L \times 0.5L$. Gaussian temperature distribution is applied to the top surface, while all the other surfaces are maintained at a lower temperature. (b) Dimensionless steady-state temperature distribution for a 3D system of size $1 \text{ mm} \times 1 \text{ mm} \times 0.5 \text{ mm}$. (c) PINN-predicted dimensionless temperature distributions along the centerline on the top surface at different length scales. (d–h) PINN-predicted contours of dimensionless steady-state temperature in the central plane ($x = 0.5L$, outlined by dashed lines in (a)) at different length scales. (i) Solution of 3D heat equation based on Fourier's law under the same boundary conditions. The temperature distribution is obtained by a well-trained PINN. Y and Z are normalized spatial coordinates. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

on a single GPU for a wide length scale range.

3. Discussion

3.1. Impact of input size on the learning performance

During the training, although we discretize the angle and frequency domains for numerical integration, the collocation points are sampled in the spatial \mathbf{x} and parameter μ spaces, without the need for mesh generation. We note that there have also been mesh-free (in spatial domain only) deterministic numerical methods for phonon BTE [66], but it was only validated for the gray model in 1D or 2D systems. As for the impact of input size on the prediction accuracy, the BTE residuals are usually evaluated on a large number of training points to ensure the learning quality. Notably, the number of collocation points in training is already smaller than those traditionally used in deterministic numerical BTE solvers [22,29,67], yet the PINN predictions remain highly accurate even for testing domains composed of many more points due to its interpolation ability.

As for the computational cost, the training time would approximately scale linearly with the total number of input points. For the conventional deterministic solvers (e.g., DOM), the computational time approximately scales linearly with the grid size in the ballistic regime [68], but they usually have very slow convergence in the diffusive regime as discussed above.

3.2. Advantages, limitations, and future perspectives of the current framework

Different from conventional methods which either solve phonon BTE iteratively on numerous nodes/cells or simulate advection and scattering of stochastic samples, the proposed PINN framework learns the high-dimensional solution function by minimizing the BTE residuals on a fair number of collocation points without using any labeled data. The advantages of this method are summarized as follows.

First, as shown in Table 1, the training cost of PINN increases with the number of dimensions, but the online evaluation can be very fast even for high-dimensional cases. Although conventional numerical methods can solve 1D and 2D problems accurately within a few minutes [32], they usually require long computation time (tens of hours) and large memory (hundreds of gigabytes) for grid-independent solutions of 3D problems even using large scale parallel computing [22,29]. However, our PINN model, while only trained on a limited number of sampled points, can provide solutions for a computational domain composed of a large number of grid points comparable to that required to achieve convergence using deterministic numerical solvers. In contrast to intricate parallel algorithms and computing, which may not be easily accessible, the proposed PINN model has the advantage of simple implementation and low evaluation cost.

More than being efficient in evaluations, the PINN framework has been shown to provide consistently accurate predictions from ballistic to diffusive regimes. By treating the characteristic length as

an input parameter, this model parameterizes the solution to phonon BTE with geometric variations, which will be beneficial for geometric optimization and uncertainty quantification. For example, if we include the geometric parameters (e.g., size of the nanopores) as inputs, we may quickly obtain the effective thermal conductivity of a nanostructured material with varying geometries using a pre-trained PINN model. In addition, PINN can also be used to propagate parameter uncertainties to the quantity of interest (e.g., heat flux, thermal conductivity) with enormous MC samples.

Furthermore, the data-free feature makes the proposed method more appealing than learning solely from the labeled data. Compared to data-driven learning with limited training data [69–71], the physics-constrained learning will be generally better since the computational domain can be sampled by an arbitrary number of collocation points. Besides, data-driven learning requires additional offline simulations and the computational cost can rapidly grow with more training points, especially when 3D problems are considered.

Although showing great promise, the current framework has several limitations, which warrants further research. First, this study is focused on parametrized steady-state phonon BTE, and the current model is not designed to handle dynamic systems. Further developments are necessary in order to solve spatiotemporal PDEs. For instance, we can employ the Long-Short Term Memory (LSTM) recurrent neural network architecture to predict the dynamics, whose effectiveness has been demonstrated in recent studies for time series prediction [72]. Second, the phonon BTE are solved under the assumption that the temperature difference is small enough such that the equilibrium state can be linearly approximated and the relaxation time is independent of space. To predict thermal transport with a large temperature difference, we need to adopt the distribution function f -based BTE with space-dependent relaxation time [31]. Third, the training cost will increase for complex problems (e.g., irregular geometries, multi-layered structures) since the equation residuals have to be calculated on a large number of collocation points in high-dimensional input spaces. We may adopt the convolutional neural network (CNN) structure [46,70,73] for complex geometries, which enables efficient learning even on irregular domains because of its parameter-sharing feature. While these directions will be explored in the future for practical applications, the current work represents the first demonstration of using PINN to efficiently solve mode-resolved phonon BTE. Our work may profoundly impact thermal transport research and make BTE a viable tool for practical device design and optimization.

4. Methods

4.1. Energy-based phonon BTE

In general, the energy-based phonon BTE [4,21,30,32,74,75], under the single mode relaxation time approximation, can be written as

$$\frac{\partial e}{\partial t} + \mathbf{v} \cdot \nabla e = \frac{e^{eq} - e}{\tau}, \quad (5)$$

where $e(\mathbf{x}, \mathbf{s}, k, p, t) = \hbar \omega D(\omega, p) [f - f^{BE}(T_{ref})]$ is the phonon energy deviational distribution function, e^{eq} is the associated equilibrium distribution function, \mathbf{v} is the group velocity, and τ is the effective relaxation time.

The phonon distribution function $f = f(\mathbf{x}, \mathbf{s}, k, p, t)$ (or $f(\mathbf{x}, \mathbf{s}, \omega, p, t)$) is a function dependent on the spatial vector \mathbf{x} , directional unit vector $\mathbf{s} = (\cos\theta, \sin\theta\cos\phi, \sin\theta\sin\phi)$ (θ is the polar angle and ϕ is the azimuthal angle), time t , wave number k (or angular frequency

$\omega = \omega(k, p)$) and polarization p . $f^{BE}(T_{ref})$ denotes the equilibrium distribution function following the Bose-Einstein distribution at the reference temperature. Assuming that the temperature difference throughout the domain is much less than the reference temperature ($|\Delta T| \ll T_{ref}$), we can make the following approximation [21,28].

$$e^{eq}(k, p, T) = \hbar \omega D(\omega, p) [f^{BE}(T) - f^{BE}(T_{ref})] \approx C(\omega, p) (T - T_{ref}), \quad (6)$$

where \hbar is the Planck's constant divided by 2π , $D(\omega, p) = \frac{k^2}{2\pi^2|\mathbf{v}|}$ is the phonon density of states, and $C(\omega, p) = \hbar \omega D(\omega, p) \frac{\partial f^{BE}}{\partial T}$ represents the modal heat capacity. This assumption has been widely used in previous studies [30,32,75], and with that we can simplify the problem significantly by using the linear relation in Eq. (6). Furthermore, the relaxation time τ and the heat capacity C can be treated as independent of space, and they are calculated based on the reference temperature T_{ref} .

In this work, we only consider the steady state. For a system without internal heat source, the first law of thermodynamics requires that the divergence of the heat flux must be zero, which can be calculated by integrating the right-hand side of Eq. (5) over the solid angle space (Ω) and frequency space (ω, p)

$$\nabla \cdot \mathbf{q} = \sum_p \int_0^{\omega_{max,p}} \int_{4\pi} \frac{e^{eq} - e}{\tau} d\Omega d\omega = 0, \quad (7)$$

where the total heat flux is

$$\mathbf{q} = \sum_p \int_0^{\omega_{max,p}} \int_{4\pi} \mathbf{v} e d\Omega d\omega, \quad (8)$$

and the local temperature can be obtained by substituting Eq. (6) into Eq. (7)

$$T = T_{ref} + \frac{1}{4\pi} \left(\sum_p \int_0^{\omega_{max,p}} \int_{4\pi} \frac{e}{\tau} d\Omega d\omega \right) \times \left(\sum_p \int_0^{\omega_{max,p}} \int_{4\pi} \frac{C}{\tau} d\Omega d\omega \right)^{-1}. \quad (9)$$

Due to non-equilibrium effects, the temperature calculated here is merely representative of the local internal energy instead of thermodynamics temperature [22,30]. For the sake of brevity, it is still referred to as the temperature in numerical experiments. Based on Eqs. (5), (7) and (8) above, we can derive Eq. (1).

4.2. Boundary conditions

Generally, three types of boundary conditions are used in phonon transport problems [30,76], including isothermal boundary condition, diffusely reflecting boundary condition and periodic boundary condition. These boundary conditions can be applied to problems with any parameter sets μ .

Isothermal boundary absorbs all phonons that strike it and emits phonons in thermal equilibrium with the boundary temperature T_b . Mathematically, this may be written as

$$e(\mathbf{x}_b, \mathbf{s}, k, p) = e^{eq}(k, p, T_b), \quad \mathbf{s} \cdot \mathbf{n}_b > 0, \quad (10)$$

where \mathbf{n}_b is the normal unit vector pointing into the simulation domain.

Diffusely reflecting boundary is a category of adiabatic boundary. At this boundary, the net heat flux is zero, and the phonons are

reflected with equal probability along all possible directions, namely,

$$e(\mathbf{x}_b, \mathbf{s}, k, p) = \frac{1}{\pi} \int_{\mathbf{s}' \cdot \mathbf{n}_b < 0} e(\mathbf{x}_b, \mathbf{s}', k, p) |\mathbf{s}' \cdot \mathbf{n}_b| d\Omega, \quad \mathbf{s} \cdot \mathbf{n}_b > 0. \quad (11)$$

When the *periodic boundary conditions* are applied, a phonon that crosses a periodic boundary is emitted at the opposite boundary with the same velocity vector and frequency. Moreover, the corresponding boundaries are subject to local thermal equilibrium, which can be expressed as

$$e(\mathbf{x}_{b_1}, \mathbf{s}, k, p) - e^{eq}(k, p, T_{b_1}) = e(\mathbf{x}_{b_2}, \mathbf{s}, k, p) - e^{eq}(k, p, T_{b_2}), \quad (12)$$

where \mathbf{x}_{b_1} , T_{b_1} and \mathbf{x}_{b_2} , T_{b_2} are the spatial coordinates and temperatures of two associated periodic boundaries b_1 and b_2 , respectively.

4.3. PINN architecture and training

The proposed PINN model is composed of two sub-DNNs for the equilibrium (e^{eq}) and non-equilibrium (e^{neq}) parts of the phonon energy distribution, respectively (Fig. 1). The sub-DNNs share the identical structure of 8 hidden layers with 30 neurons per layer. Two sub-DNNs are trained simultaneously with a unified physics-informed loss function. We employ the Swish activation function ($x \cdot \text{Sigmoid}(x)$) [77] in each layer except the last one, where a linear activation function is applied. The Adam optimizer [78], a robust variant of the stochastic gradient descent algorithm, is used to solve the optimization problem defined in Eq. (3) by randomly sampling mini-batches of inputs (see Fig. S4 for an example of training process). The initial learning rate is set as 1×10^{-3} , and training points are generated by discretization of the input domain. To approximate the integrals in Eq. (1), the Gauss-Legendre quadrature [79] is adopted for the solid angle space, while the midpoint rule is used for the frequency space. In the case the spatial domain is logically rectangular, we can set the interior training points as low-discrepancy Sobol sequences [80] to alleviate the curse of dimensionality (Supplementary Note 3). Input spatial coordinates are scaled to the range [0, 1]. The output for the non-equilibrium part e^{neq} is scaled with Kn to make it of the order $O(1)$. The PINN algorithm is implemented within the PyTorch platform [81], and all numerical experiments are performed on a single NVIDIA GeForce TITAN Xp Graphic Processing Unit (GPU).

Credit author statement

Ruiyang Li: Conceptualization, Investigation, Coding, Writing – original draft. Eungkyu Lee: Conceptualization, Writing – reviewing & editing. Tengfei Luo: Conceptualization, Supervision, Writing – reviewing & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank ONR MURI (N00014-18-1-2429) for the financial support. The simulations are supported by the Notre Dame Center for Research Computing, and NSF through the

eXtreme Science and Engineering Discovery Environment (XSEDE) computing resources provided by Texas Advanced Computing Center (TACC) Stampede II under grant number TG-CTS100078. This work is also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1C1C1006251).

Data availability

The code and data required to reproduce these findings are available to download from <https://github.com/RuiyangLi6/PINN-pBTE> upon publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mtphys.2021.100429>.

References

- [1] E.S. Toberer, L.L. Baranowski, C. Dames, *Annu. Rev. Mater. Res.* 42 (2012) 179.
- [2] Y. Guo, M. Wang, *Phys. Rep.* 595 (2015) 1.
- [3] A. Minnich, *J. Phys. Condens. Matter* 27 (2015): 053202.
- [4] G. Chen, *Nanoscale Energy Transport and Conversion: A Parallel Treatment of Electrons, Molecules, Phonons, and Photons*, Oxford university press, 2005.
- [5] Z.M. Zhang, Z.M. Zhang, Luby, *Nano/microscale Heat Transfer*, Springer, 2007.
- [6] G. Chen, *Phys. Rev. B* 57 (1998) 14958.
- [7] Y.-C. Hua, B.-Y. Cao, *Int. J. Heat Mass Tran.* 78 (2014) 755.
- [8] S. Hamian, T. Yamada, M. Faghri, K. Park, *Int. J. Heat Mass Tran.* 80 (2015) 781.
- [9] C. Zhang, Z. Guo, S. Chen, *Phys. Rev.* 96 (2017): 063311.
- [10] Y. Ju, K. Goodson, *Appl. Phys. Lett.* 74 (1999) 3005.
- [11] J.Y. Murthy, S.V. Narumanchi, P.-G. Jose'A, T. Wang, C. Ni, S.R. Mathur, *Int. J. Multiscale Comput. Eng.* 3 (2005).
- [12] S. Mazumder, A. Majumdar, *J. Heat Tran.* 123 (2001) 749.
- [13] D. Lacroix, K. Joulain, D. Lemonnier, *Phys. Rev. B* 72 (2005): 064305.
- [14] A. Mittal, S. Mazumder, *J. Heat Tran.* 132 (2010).
- [15] J.-P.M. Péraud, N.G. Hadjiconstantinou, *Phys. Rev. B* 84 (2011) 205331.
- [16] Z. Shomali, B. Pedar, J. Ghazanfarian, A. Abbassi, *Int. J. Therm. Sci.* 114 (2017) 139.
- [17] R. A. Escobar and C. H. Amon, (2007).
- [18] R.A. Escobar, C.H. Amon, *J. Heat Tran.* 130 (2008).
- [19] J. Murthy, S. Mathur, *J. Heat Tran.* 124 (2002) 1176.
- [20] S.V. Narumanchi, J.Y. Murthy, C.H. Amon, *J. Heat Tran.* 126 (2004) 946.
- [21] A.J. Minnich, G. Chen, S. Mansoor, B. Yilbas, *Phys. Rev. B* 84 (2011) 235207.
- [22] S.A. Ali, G. Kollu, S. Mazumder, P. Sadayappan, A. Mittal, *Int. J. Therm. Sci.* 86 (2014) 341.
- [23] J.-P.M. Péraud, N.G. Hadjiconstantinou, *Phys. Rev. B* 91 (2015) 235321.
- [24] A. Chattopadhyay, A. Pattamatta, *Numer. Heat Tran., Part B: Fundamentals* 66 (2014) 360.
- [25] Y. Guo, M. Wang, *J. Comput. Phys.* 315 (2016) 1.
- [26] M.L. Adams, E.W. Larsen, *Prog. Nucl. Energy* 40 (2002) 3.
- [27] J.M. Loy, S.R. Mathur, J.Y. Murthy, *J. Heat Tran.* 137 (2015).
- [28] J.M. Loy, J.Y. Murthy, D. Singh, *J. Heat Tran.* 135 (2013).
- [29] P. Allu, S. Mazumder, *Int. J. Heat Mass Tran.* 100 (2016) 165.
- [30] X.-P. Luo, H.-L. Yi, *Int. J. Heat Mass Tran.* 114 (2017) 970.
- [31] C. Zhang, Z. Guo, *Int. J. Heat Mass Tran.* 134 (2019) 1127.
- [32] C. Zhang, Z. Guo, S. Chen, *Int. J. Heat Mass Tran.* 130 (2019) 1366.
- [33] X. Qian, S. Peng, X. Li, Y. Wei, R. Yang, *Mater. Today Phys.* 10 (2019) 100140.
- [34] R. Li, E. Lee, T. Luo, *Mater. Today Phys.* 12 (2020) 100181.
- [35] R. Li, Z. Liu, A. Rohskopf, K. Gordiz, A. Henry, E. Lee, T. Luo, *Appl. Phys. Lett.* 117 (2020) 152102.
- [36] Y.-B. Liu, J.-Y. Yang, G.-M. Xin, L.-H. Liu, G. Csányi, B.-Y. Cao, *J. Chem. Phys.* 153 (2020) 144501.
- [37] Y. LeCun, Y. Bengio, G. Hinton, *Nature* 521 (2015) 436.
- [38] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, *Deep Learning*, vol. 1, MIT press Cambridge, 2016, p. 2.
- [39] F. Scarselli, A.C. Tsoi, *Neural Network*. 11 (1998) 15.
- [40] I.E. Lagaris, A. Likas, D.I. Fotiadis, *IEEE Trans. Neural Network*. 9 (1998) 987.
- [41] J. Berg, K. Nyström, *Neurocomputing* 317 (2018) 28.
- [42] J. Sirignano, K. Spiliopoulos, *J. Comput. Phys.* 375 (2018) 1339.
- [43] M. Raissi, P. Perdikaris, G.E. Karniadakis, *J. Comput. Phys.* 378 (2019) 686.
- [44] J. Han, A. Jentzen, E. Weinan, *Proc. Natl. Acad. Sci. Unit. States Am.* 115 (2018) 8505.
- [45] L. Yang, D. Zhang, and G. E. Karniadakis, *arXiv preprint arXiv:1811.02033* (2018).
- [46] Y. Zhu, N. Zabarar, P.-S. Koutsourelakis, P. Perdikaris, *J. Comput. Phys.* 394 (2019) 56.
- [47] Q. Lou, X. Meng, and G. E. Karniadakis, (2020), p. arXiv:2010.09147.
- [48] L. Sun, H. Gao, S. Pan, J.-X. Wang, *Comput. Methods Appl. Mech. Eng.* 361

- (2020) 112732.
- [49] L. Lu, X. Meng, Z. Mao, G.E. Karniadakis, *SIAM Rev.* 63 (2021) 208.
- [50] A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind, *J. Mach. Learn. Res.* 18 (2017) 5595.
- [51] P. Grohs, F. Hornung, A. Jentzen, and P. Von Wurstemberger, arXiv preprint arXiv:1809.02362 (2018).
- [52] M. Hutzenthaler, A. Jentzen, T. Kruse, T.A. Nguyen, *SN Part. Diff. Eq. Appl.* 1 (2020) 1.
- [53] S. Karumuri, R. Tripathy, I. Bilonis, J. Panchal, *J. Comput. Phys.* 404 (2020) 109120.
- [54] K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of the IEEE International Conference on Computer Vision*(2015), pp. 1026.
- [55] M. Holland, *Phys. Rev.* 132 (1963) 2461.
- [56] E. Pop, R.W. Dutton, K.E. Goodson, *J. Appl. Phys.* 96 (2004) 4998.
- [57] D. Terris, K. Joulain, D. Lemonnier, D. Lacroix, *J. Appl. Phys.* 105 (2009): 073516.
- [58] B. Liao, B. Qiu, J. Zhou, S. Huberman, K. Esfarjani, G. Chen, *Phys. Rev. Lett.* 114 (2015) 115901.
- [59] C.J. Glassbrenner, G.A. Slack, *Phys. Rev.* 134 (1964) A1058.
- [60] D.G. Cahill, H.E. Fischer, T. Klitsner, E. Swartz, R. Pohl, *J. Vac. Sci. Technol.: Vac. Surf. Films* 7 (1989) 1259.
- [61] D.G. Cahill, M. Katiyar, J. Abelson, *Phys. Rev. B* 50 (1994) 6077.
- [62] D.P. Sellan, J. Turney, A.J. McGaughey, C.H. Amon, *J. Appl. Phys.* 108 (2010) 113524.
- [63] C. Hua, A.J. Minnich, *J. Appl. Phys.* 117 (2015) 175306.
- [64] J.C. Chai, H.S. Lee, S.V. Patankar, *Numer. Heat Tran., Part B Fundamentals* 24 (1993) 373.
- [65] J. Cuffe, et al., *Phys. Rev. B* 91 (2015) 245423.
- [66] S. Zahiri, C. Shao, Y. Shen, H. Bao, *Numer. Heat Tran., Part B: Fundamentals* 70 (2016) 459.
- [67] S. Zahiri, Z. Xu, Y. Hu, H. Bao, Y. Shen, *Int. J. Heat Mass Tran.* 138 (2019) 267.
- [68] V. Fiveland, J. Jessee, *J. Thermophys. Heat Tran.* 10 (1996) 445.
- [69] H. Wei, S. Zhao, Q. Rong, H. Bao, *Int. J. Heat Mass Tran.* 127 (2018) 908.
- [70] B. Kim, V.C. Azevedo, N. Thuerey, T. Kim, M. Gross, B. Solenthaler, *Computer Graphics Forum, Wiley Online Library*, 2019, p. 59.
- [71] Q. Rong, H. Wei, X. Huang, H. Bao, *Compos. Sci. Technol.* 184 (2019) 107861.
- [72] R. Zhang, Y. Liu, H. Sun, *Comput. Methods Appl. Mech. Eng.* 369 (2020) 113226.
- [73] H. Gao, L. Sun, J.-X. Wang, *J. Comput. Phys.* 428 (2021) 110079.
- [74] C. Hua, A.J. Minnich, *Phys. Rev. B* 89 (2014): 094302.
- [75] C. Zhang, S. Chen, Z. Guo, and L. Wu, arXiv preprint arXiv:1812.06405 (2018).
- [76] Z. Guo, K. Xu, *Int. J. Heat Mass Tran.* 102 (2016) 944.
- [77] P. Ramachandran, B. Zoph, and Q. V. Le, arXiv preprint arXiv:1710.05941 (2017).
- [78] D. P. Kingma and J. Ba, arXiv preprint arXiv:1412.6980 (2014).
- [79] N. Hale, A. Townsend, *SIAM J. Sci. Comput.* 35 (2013) A652.
- [80] I.y.M. Sobol, *Zh. Vychislitel noi Mat. Mat. Fiz.* 7 (1967) 784.
- [81] A. Paszke, et al., *NIPS Workshop*, 2017.

Supplementary Information for

Physics-Informed Neural Networks for Solving Multiscale Mode-Resolved Phonon Boltzmann Transport Equation

Ruiyang Li, Eungkyu Lee*, and Tengfei Luo*

* Corresponding author. Email: eleest@khu.ac.kr (E. Lee); tluo@nd.edu (T. Luo)

Supplementary Note 1: Phonon dispersion and scattering

The dispersion relations of the acoustic phonons are approximated as [1]

$$\omega = c_1 k + c_2 k^2,$$

where for LA branch $c_1 = 9.01 \times 10^5$ cm/s, $c_2 = -2.0 \times 10^{-3}$ cm²/s; for TA branch $c_1 = 5.23 \times 10^5$ cm/s, $c_2 = -2.26 \times 10^{-3}$ cm²/s. The Matthiessen's rule is used to estimate the effective relaxation time by combining different scattering processes [2], including the impurity scattering, umklapp (U) and normal (N) phonon-phonon scattering,

$$\tau^{-1} = \tau_{impurity}^{-1} + \tau_U^{-1} + \tau_N^{-1} = \tau_{impurity}^{-1} + \tau_{NU}^{-1},$$

where the relaxation time formulas and coefficients are given in Table S1. We note that the dispersion and relaxation times can also be from first-principles calculations for each discrete mode [3]. The PINN implementation will not change.

Table S1 Relaxation time formulas and coefficients [4].

$\tau_{impurity}^{-1}$	$A_i \omega^4, A_i = 1.498 \times 10^{-45} \text{ s}^3$
LA	$\tau_{NU}^{-1} = B_L \omega^2 T^3, B_L = 1.180 \times 10^{-24} \text{ K}^{-3}$
TA	$\tau_{NU}^{-1} = B_T \omega T^4, 0 \leq k < \pi/a$
	$\tau_{NU}^{-1} = B_U \omega^2 / \sinh(\hbar \omega / k_B T), \pi/a \leq k \leq 2\pi/a$
	$B_T = 8.708 \times 10^{-13} \text{ K}^{-3}, B_U = 2.890 \times 10^{-18} \text{ s}$
	$a = 0.5431 \text{ nm}$

Supplementary Note 2: Convergence test of N_x

For the training process, a large N_x is preferred such that the model can accurately learn the solution function in the whole spatial domain, but the training time will increase with a larger N_x . In the 1D case, we have trained the PINNs with different $N_{x,train}$ and tested the with the same setting ($N_{x,test} = 100$). Figure S1(a) shows the prediction errors (compared to the analytical solutions) as a function of $N_{x,train}$. It is observed that the mean relative errors in both T^* and k_{eff}/k_{bulk} converge when $N_{x,train} > 30$, which implies that it is sufficient to set $N_{x,train} = 40$ during the training.

For the testing process, N_x does not show significant impact on the accuracy if it is sufficiently large especially when evaluating the heat flux or effective thermal conductivity. We have tested the model trained with $N_{x,train} = 40$ by using different $N_{x,test}$ in the testing processes. As shown in Fig. S1(b), the error in T^* increases with $N_{x,test}$, because more and more unseen points are evaluated in the testing when $N_{x,test} > N_{x,train}$. However, the error tends to stay at around 1% as $N_{x,test}$ continues to go up. On the other hand, the error in k_{eff}/k_{bulk} is very stable with respect to $N_{x,test}$.

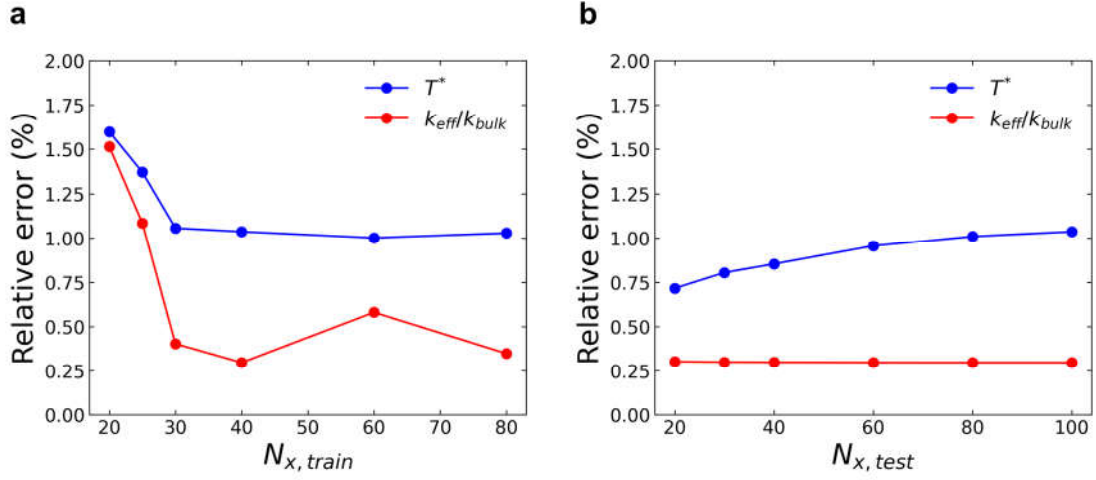


Fig. S1. Relative test error for the 1D case with varying (a) $N_{x,train}$ ($N_{x,test} = 100$), or with varying (b) $N_{x,test}$ ($N_{x,train} = 40$).

Supplementary Note 3: Training points in the spatial domain

We use low-discrepancy Sobol sequences [5] for logically rectangular spatial domains to constitute the interior training sets. Sobol sequences arise in the context of Quasi-Monte Carlo integration [6], which can alleviate the curse of dimensionality and provide more flexibility for high-dimensional problems. Figure S2 shows the spatial domains containing interior training points in 2D phonon transport problems. In 2D in-plane cases, we use 300 Sobol points for the interior domain, and 30 equidistant points for each boundary. In 2D square cases, for better description of phonon transport, we construct a non-uniform spatial domain (450 Sobol points) with more training points close to the hot boundary.

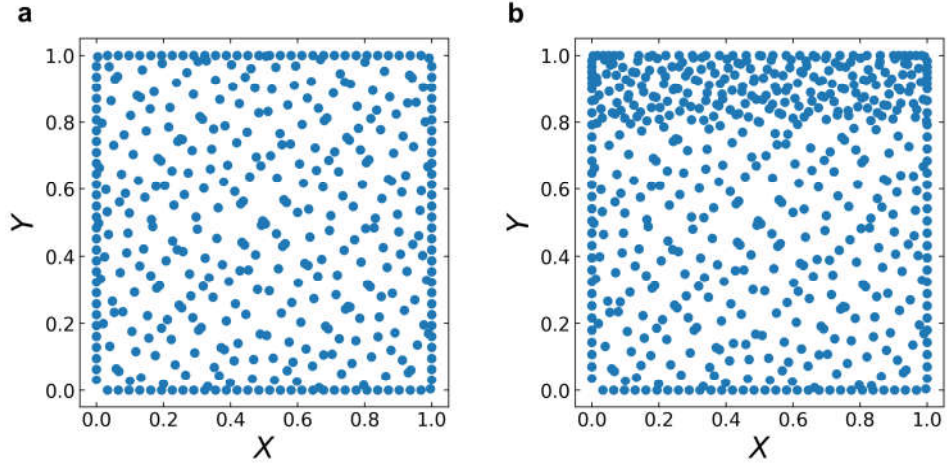


Fig. S2. Spatial training points based on Sobol sequences for (a) the 2D in-plane case, and (b) the 2D square case.

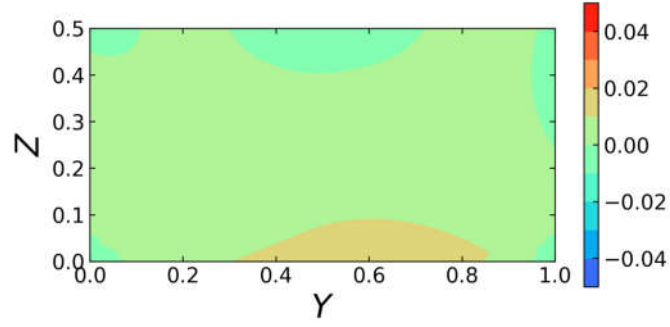


Fig. S3. Difference between PINN predicted temperature contour (in a 3D system of size $1\text{ mm} \times 1\text{ mm} \times 0.5\text{ mm}$) and solution of 3D heat equation based on Fourier's law in the central plane. The error rate is less than 1.4 %.

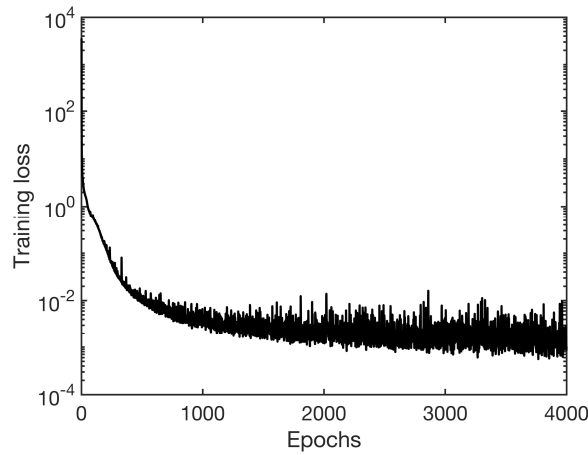


Fig. S4. Example of learning curve in 2D in-plane phonon transport problem. The training loss is averaged among 15 mini-batches for each epoch.

Supplementary references

- [1] E. Pop, R. W. Dutton, and K. E. Goodson, *Journal of Applied Physics* **96**, 4998 (2004).
- [2] J. Y. Murthy, S. V. Narumanchi, P.-G. Jose'A, T. Wang, C. Ni, and S. R. Mathur, *International Journal for Multiscale Computational Engineering* **3** (2005).
- [3] K. Esfarjani, G. Chen, and H. T. Stokes, *Physical Review B* **84**, 085204 (2011).
- [4] D. Terris, K. Joulain, D. Lemonnier, and D. Lacroix, *Journal of Applied Physics* **105**, 073516 (2009).
- [5] I. y. M. Sobol', *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* **7**, 784 (1967).
- [6] R. E. Caflisch, *Acta numerica* **1998**, 1 (1998).