CHAPTER 3

MACHINE LEARNING FOR THERMAL PROPERTIES OF INORGANIC CRYSTALS

Ogheneyoma Aghoghovbia, Joshua Ojih, & Ming Hu*

Department of Mechanical Engineering, University of South Carolina, Columbia, SC

*Address all correspondence to: Ming Hu, E-mail: HU@sc.edu

This article reviews artificial intelligence (AI) and machine learning (ML) approaches for predicting transport behaviors of thermal energy carriers, mainly lattice vibrations (phonons), in inorganic crystals. Traditional first principles approaches, namely density functional theory (DFT) combined with the Boltzmann transport equation (BTE) (DFT+BTE), provide accurate predictions but are computationally expensive and not feasible for handling large-scale screening of unknown materials. State-of-the-art AI/ML methodologies provide computationally efficient solutions by learning intricate structure-property relations, allowing for fast calculations of phonon-related thermal properties at minimal computational expense. We classify AI/ML methodologies into two groups: direct method, which directly predicts properties like lattice thermal conductivity (κ_L) from structural descriptors or graphs, and indirect method, which substitutes computationally demanding force evaluations in DFT calculations with ML-predicted forces or interatomic force constants, thus retaining physical interpretability. The recent developments in graph neural networks, machine learning interatomic potentials, and data-driven feature engineering are discussed, highlighting their scalability, accuracy, and applicability to a wide chemical space. The article also outlines current challenges, including dataset quality, transferability, and integration with high-throughput computational frameworks, and identifies future directions for AI/ML accelerated thermal materials discovery.

KEY WORDS: *phonons, thermal transport, machine learning*

1. INTRODUCTION

Phonon (lattice vibration) and electron transport behavior influence the thermal properties of inorganic crystal structures. Phonons are quasi-particles or quantized lattice excitations that describe the collective motion of atoms. They not only carry atomic vibrational energy while propagating through the crystal lattice but also act as the primary carriers of heat energy for semiconductors and insulators. According to the kinetic theory, the lattice thermal conductivity (κ_L) of a crystalline material is expressed by $\kappa_L = (1/3)C_v V_g \bar{l}$, where C_v is the constant volume heat capacity, V_g is the average phonon group velocity, and $\bar{l}$ is the average phonon mean free path (MFP).[1]

In semiconductors and insulators, phonons are the dominant contributors to thermal conductivity, and their transport mechanisms are influenced by both lattice harmonicity and anharmonicity, represented by the intrinsic phonon-phonon interactions. Phonon transport in solids is also influenced by extrinsic phonon scatterings, such as defects,

ISSN: 1049–0787; ISBN: 978–1–56700–554–7/25/$35.00 + $00.00

dimensionality [one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D)], length, interface, and grain boundaries.[2] Important indicators for describing phonon transport properties from the perspective of harmonicity and anharmonicity include phonon mode, defined by frequency and wave-vector, phonon dispersions, specific heat, phonon density of states (DOS), group velocity, scattering rate, MFP, and lifetime.[1] Strong lattice anharmonicity causes substantial phonon-phonon scattering in materials with intricate crystal structures, frequently lowering thermal conductivity.[2]

Electrons, particularly in metals and metallic systems, also influence thermal conductivity. Heat flow is facilitated by the electrons' unrestricted movement inside the lattice, which transfers energy when they collide with the lattice ions.[3,4] Electron density and scattering mechanisms determine the electron contribution to thermal conductivity, which is essential in metals. Electron scattering brought on by lattice defects can increase or decrease thermal conductivity.[5] The interaction of phonons and electrons gives inorganic crystals their thermal characteristics. Temperature and crystal structure are two variables affecting these carriers' respective contributions. In materials with large electron and phonon contributions, improving electron mobility while decreasing phonon scattering is necessary to maximize thermal conductivity.[6] Many cutting-edge technological applications rely on manipulating materials' phonon transport, including thermoelectrics,[7,8] superconductors,[9,10] energy conversion,[11] thermal management,[12] quantum computing,[13,14] and photovoltaics.[15,16] For instance, in order to maximize the dimensionless figure of merit (ZT), a measure used to evaluate the energy conversion efficiency of thermoelectric materials,[17] phonon transport must be suppressed to ensure low thermal conductivity while maintaining high electrical conductivity. High ZT values indicate better thermoelectric performance or the material's ability to convert heat into electricity more effectively.[18,19] Similarly, minimizing phonon scattering is also vital in superconductors because electron pairing, which is essential to the phenomenon of superconductivity, is facilitated by lattice vibrations through the electron-phonon coupling mechanism.[20]

To characterize phonon transport in materials, experimental techniques such as X-ray diffraction,[21–23] infrared spectroscopy,[24–26] Raman spectroscopy,[27,28] inelastic neutron scattering,[29,30] and thermoreflectance thermal imaging[31] have been used to measure phonon transport properties but are limited in several ways. Isolating phonon contribution from other heat carriers like electrons requires high precision, and techniques like Raman spectroscopy and inelastic neutron scattering require expensive experimental setups and are also limited in their application because they frequently require large, high-quality single crystals.[32] It is important to note that experimental measurements are susceptible to external influences. Impurities, defects, and boundary conditions make it difficult for experimentalists to obtain repeatable results.

Because of the complex structure of atomic interactions and lattice dynamics in materials, phonon properties' prediction is computationally fraught with difficulties.[33] Both harmonic and anharmonic interactions, which control phonon dispersion relations and scattering mechanisms, respectively, must be thoroughly understood in order to simulate these features accurately.[34] It is challenging to create predictive models that can thoroughly represent phonon transport activity across many materials due to the inherent complexity

of these interactions. Traditional methods, including density functional theory (DFT), coupled with the Boltzmann transport equation (BTE) calculations[35–37] and molecular dynamics (MD) simulations,[37,38,40] have been utilized to predict phonon transport properties. Although the DFT+BTE approach accurately predicts the phonon transport properties without any arbitrary input parameters, it is computationally expensive and one must consider different levels of phonon-phonon scattering explicitly, such as the three-phonon process and four-phonon process. MD, on the other hand, naturally captures all orders of anharmonic effect and temperature dependency but lacks quantum accuracy, and the quality of an MD simulation depends largely on the availability and accuracy of interatomic potentials. Resolving mode resolved phonon transport information is another challenge for MD simulation of phonon transport, which usually would require separate post-processing of MD simulation results. Both methods face grand challenges when dealing with large-scale unknown materials, in terms of computational efficiency (for DFT+BTE) and accuracy (for MD).

Artificial intelligence (AI) and machine learning (ML) are promising substitutes for these conventional methods. On the one hand, AI/ML models can be trained to map highly nonlinear relationships between atomic structures and material's phonon transport properties, and such models are usually called the *direct* method. The AI/ML models' flexibility enables them to efficiently handle the complexity and diversity inherent in phonon property predictions, thereby overcoming DFT+BTE limitations,[41] and therefore can considerably decrease the computational requirements for phonon transport property predictions. On the other hand, in the phonon transport workflow, AI/ML algorithms, such as machine learning potentials (MLPs), can replace the computationally expensive DFT calculations by producing predictions orders of magnitude faster than conventional techniques with comparable accuracy. Such an approach is called the *indirect* method. This makes it possible to quickly assess material properties across large design spaces, which is essential for the practical discovery of new materials.[37] AI/ML-assisted MLPs can also replace the critical part of the interatomic potentials in the MD simulations and thus can accelerate the prediction of phonon transport properties of many new materials.

A few review papers have discussed the use of ML in materials science, with emphasis on property prediction and materials discovery. Most reviews, however, either present a broad overview of ML approaches across diverse materials properties,[42–44] or focus on specific applications such as bandgap prediction,[45] formation energy,[46] or mechanical properties.[47] Note that some papers have reviewed progress of ML applications in heat transfer, which usually focus on continuum-level thermal transport research[48,49] or a limited number of phonon transport studies using traditional ML.[48,50,51] In contrast, our review is dedicated to the comprehensive applications of ML techniques (both traditional and advanced ML algorithms) for predicting phonon transport in inorganic crystals, where accurate modeling remains computationally demanding. We provide a twofold perspective: (i) direct models, in which ML directly maps crystal structures to thermal properties, and (ii) indirect models, in which ML accelerates conventional workflows by predicting atomic forces or interatomic force constants (IFCs). This framework structures the core of the review and allows us to synthesize developments in both descriptor-based and graphic neural

network (GNN)-based methods, as well as ML-interatomic potentials. We further examine model generalization, data scarcity, and inverse design, which are rarely emphasized in earlier reviews. Together, these elements position our work as a timely and distinctive contribution to the field.

As illustrated in the schematic overview presented in Fig. 1, this review will first examine the traditional computational methods for phonon transport properties prediction and their inherent drawbacks, in particular from a high-throughput screening point of view. We will then review the later developments in AI/ML models as an emerging alternative in predicting the phonon transport properties, offering promising solutions to traditional approaches' inherent limitations in terms of computational cost, scalability, and accuracy, paving the way for efficient properties prediction, thus accelerating the discovery of novel materials with tailored thermal properties.

2. BRIEF SURVEY OF MAJOR TRADITIONAL COMPUTATIONAL METHODS FOR PHONON TRANSPORT IN INORGANIC CRYSTALS

2.1 DFT+BTE Approach

DFT is a quantum mechanical computational technique used in physics and chemistry to investigate the electronic structure of a many-body system. It is based on the first principle (*ab initio*) calculations. It is an effective method for ascertaining ground-state properties by using electron density instead of wavefunctions, thus reducing the computing complexities relative to the conventional many-body approach, which involves using

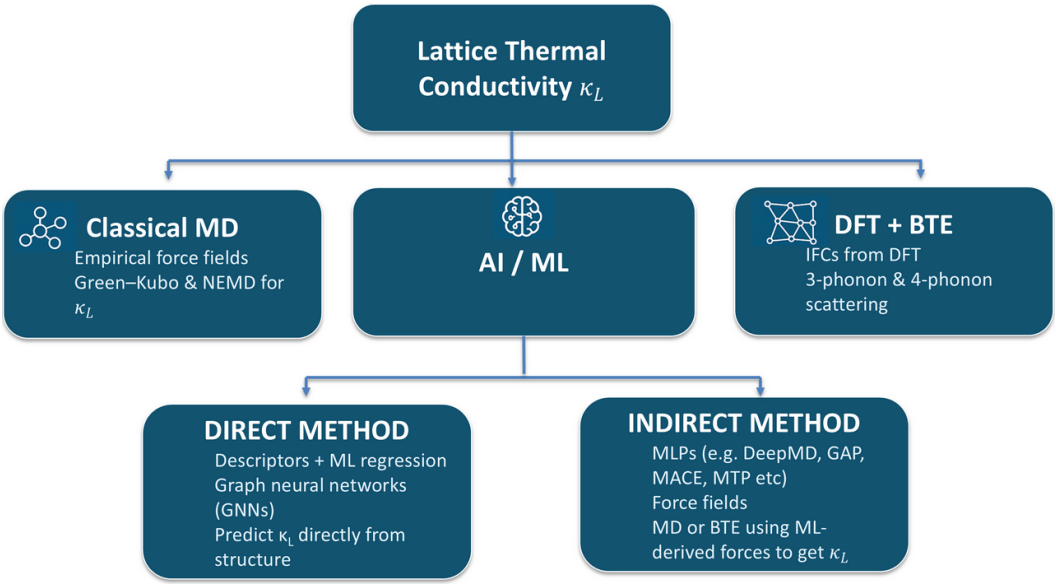

FIG. 1: Overview of the different approaches for predicting lattice thermal conductivity (κ_L), including first principles methods and machine learning–based strategies

Born–Oppenheimer approximations to solve the Schrödinger equation. The interaction of the electrons with each other gives rise to the many-body problem $\hat{H}\psi = E\psi$, where $\hat{H}$ is Hamiltonian operator, ψ is the wave functional, and E is the energy of the system. The Hohenberg–Kohn theorems provide the foundation for DFT, specifically, the first states that the ground-state properties of a many-electron system are uniquely determined by the electron density $n(r)$, However, by minimizing the energy functionals $E[n(r)]$, the second theorem offers a variational principle for determining the ground-state electron density.[52] These theorems reduce the dependency from the wavefunctions to the electron density, thus simplifying the original many-body problem. The energy function is expressed as follows:[52,53]

$$E[n(r)] = Te[n(r)] + V_{\text{ext}}[n(r)] + VH[n(r)] + E_{\text{xc}}[n(r)] \tag{1}$$

where $Te[n(r)]$ is the kinetic energy of a noninteracting electron gas; $V_{\text{ext}}[n(r)]$ is the external potential energy due to nuclei; $VH[n(r)]$ is the Hartree (Coulomb) interaction energy; and $E_{\text{xc}}[n(r)]$ is the exchange-correlation (xc) energy. Kohn and Sham proposed a method that reduces the complex interacting electron systems to a simpler, noninteracting system subject to an effective potential in order to make the Hohenberg–Kohn formalism computationally practical. The Kohn–Sham equation is expressed as follows:[53]

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{KS}}(r)\right]\psi_i(r) = \epsilon_i\psi_i(r) \tag{2}$$

The effective potential V_{KS} is defined as follows:

$$V_{\text{KS}}(r) = V_{\text{ext}}(r) + V_{\text{H}}(r) + V_{\text{xc}}(r) \tag{3}$$

where $-(h^2/2m)\nabla^2$ represents the kinetic energy operator for noninteracting electrons and V_{xc} is the exchange-correlation potential. The key to the Kohn–Sham framework is the exchange-correlation function, which captures all quantum-mechanical exchange and correlation interactions among electrons. Although the exact form of V_{xc} remains unknown,[53] various approximations have been developed to enable practical DFT calculations, including the local density approximation, the generalized gradient approximations,[54] hybrid functionals, such as the Perdew–Burke–Ernzehof with zero empirical parameters (PBE0),[55] and many others. The electron density $n(r)$ form the occupied Kohn–Sham orbitals $\psi_i(r)$ is given by[53]

$$n(r) = \sum_{i=1}^{N} |\psi_i(r)|^2 \tag{4}$$

where N is the number of electrons in the system.

Calculating the systems' total energy and atomic forces is made possible by the ground-state electron density that results from choosing the proper exchange-correlation functional and solving the Kohn–Sham equations self-consistently.[53,56] Structural equilibria, i.e., the equilibrium positions of atoms, are determined using the ground-state forces.[57] Derivatives of the total energy with respect to atomic displacements around the equilibrium positions are evaluated in order to calculate the IFCs.[58,59] These IFCs at different orders are fundamental input for lattice dynamics analysis, such as phonon dispersion curves, phonon

lifetimes, and thermal conductivity computations within harmonic and anharmonic frameworks.[60–62] From this process, one can see that the atomic forces or IFCs are the foundation of phonon property prediction.

The finite difference method (FDM) is one of the major methods used to calculate the IFCs. In the finite difference approach, the atoms are displaced from their equilibrium positions and the resulting atomic forces are calculated using DFT. The number of required atoms' movements depends on the complexity and symmetry of the material. Generally speaking, the more complex and lower symmetry the material is, the more atoms' movement is required. This is followed by numerical differentiation of the atomic forces with respect to the displacements to determine the IFCs. The second-order force constant is given by[58]

$$\Phi_{I\alpha,J\beta} = -\frac{\partial F_{I\alpha}}{\partial u_{J\beta}} \tag{5}$$

where $F_{I\alpha}$ is the force acting on atom I in direction α due to displacement $u_{J\beta}$ of atom J in direction β.[58] The second-order IFCs are often calculated using software packages such as PHONOPY,[58] in conjunction with DFT codes like VASP,[63,64] Quantum ESPRESSO,[65] and ABINIT.[66,67] Third-order IFCs required for phonon scattering and thermal conductivity are usually calculated using additional software tools, such as ShengBTE[62] and PHONO3PY,[68] in addition to DFT computations using VASP or Quantum ESPRESSO.

For complex inorganic crystals, such as those with a large number of atoms in the primitive cells and/or with low material symmetries, FDM is not very efficient to get IFCs because the number of DFT runs required to get those force derivatives through FDM is huge and thus it is computationally unbearable. To solve such an issue, methods like compressive sensing lattice dynamics (CSLD) are developed, which require fewer displacement-force calculations than the traditional FDM and thus reduce computational cost. In contrast to FDM, all-order IFCs can be efficiently calculated using CSLD,[69,70] which exploits the inherent sparsity of IFC tensors. Instead of relying on extensive DFT evaluations involving many times of atoms' movement, CSLD uses compressive sensing optimization to reconstruct the complete IFC tensor after selecting a small number of atomic displacements. By reducing the amount of DFT computations required, CSLD technique drastically lowers computational costs. Accurately obtaining the IFCs is crucial for solving the BTE, a statistical framework used to predict phonon transport properties and to calculate κ_L. The BTE links the microscopic properties of phonons, such as their velocities, lifetimes, and scattering events, to the macroscopic observable thermal transport behavior, particularly κ_L.[71,72] The κ_L under the relaxation time approximation (RTA) is expressed as follows:[68]

$$\kappa_L^{\alpha\beta} = \frac{1}{V}\sum_{\lambda} C_\lambda v_\lambda^\alpha v_\lambda^\beta \tau_\lambda \tag{6}$$

where the mode-specific heat capacity C_λ is given by

$$C_\lambda = \frac{1}{Nk_B T^2} n_\lambda^0 \left(n_\lambda^0 + 1\right) (\hbar\omega_\lambda)^2 \tag{7}$$

where α and β are the cartesian coordinate directions (x, y, z), V is the volume of the unit cell, $C_\lambda v_\lambda^\alpha$ is the component of the phonon group velocity for the phonon mode λ, τ_λ is the phonon relaxation time for the phonon mode λ, k_B is the Boltzmann constant, T is the absolute temperature, N is the total number of phonon wave vectors, $\hbar$ is the reduced Planck constant, ω_λ is the phonon frequency, and n_λ^0 is the phonon occupation number and follows the Bose–Einstein distributions expressed as follows:[68]

$$n_\lambda^0 = \frac{1}{e^{\hbar\omega_\lambda/k_B T} - 1} \tag{8}$$

Computational tools, such as the ShengBTE package, which solves the phonon BTE using second- and third-order IFCs, have made substantial progress in accurately calculating κ_L from first principles. The ShengBTE software computes phonon dispersions, lifetimes, and κ_L by explicitly considering phonon scattering processes.[62] Figure 2 illustrates the DFT+BTE workflow for calculating κ_L, showing the sequence from the atomic structure input to the evaluation of phonon properties and solution of phonon BTE.

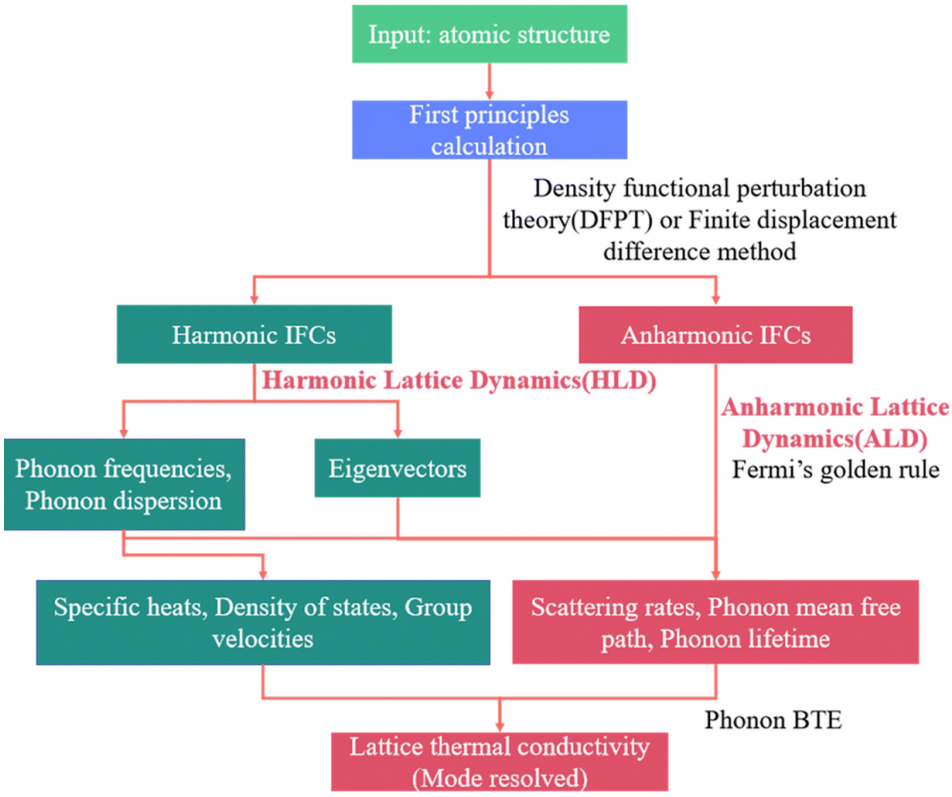

FIG. 2: Schematic workflow of the DFT+BTE approach for calculating the mode-resolved lattice thermal conductivity (κ_L) (Reprinted with permission from the Royal Society of Chemistry, Copyright 2025)[1]

The DFT+BTE approach for predicting phonon properties of inorganic crystals have been extensively used in the past decade (2015–2025).[73–77] Lindsay et al. used first principles DFT combined with an iterative solution to the linearized phonon BTE to predict cubic boron arsenide (BAs) with an exceptionally high κ_L exceeding 2×10^3 $\mathrm{Wm^{-1}K^{-1}}$ at room temperature, in rivalry to that of diamond and graphite, well known for their superior thermal conductivities.[78] As shown in Eq. (6), thermal conductivity depends on the heat capacity of phonon modes, group velocity, and lifetime. The work by Lindsay et al.[78] highlighted that the extraordinarily long lifetimes of phonons in BAs, particularly for the heat-carrying acoustic modes, are the result of suppressed three-phonon scattering rates stemming from the large bandgaps $\sim$ 10 THz between the acoustic and optical branches, which effectively limits the phase space for phonon-phonon scattering, thus leading to higher thermal conductivity. Three-phonon scattering describes the fundamental anharmonic interactions between phonons, where a single phonon decays into two others, $\omega_\lambda = \omega_{\lambda'} + \omega_{\lambda''}$, or two phonons combine to form a third $\omega_\lambda + \omega_{\lambda'} = \omega_{\lambda''}$. These interactions limit phonon lifetimes and are important in determining the κ_L.[79]

By comparing BAs with other III-V boron compounds, Lindsay et al.[78] demonstrated that while other materials, such as BP and BSb, also possess high thermal conductivity. BAs exhibits high group velocity and extremely low scattering, making it applicable for thermal management.[78,80,81] While BAs demonstrates exceptionally high κ_L, attributed to its intrinsically weak anharmonicity and the suppression of three-phonon scattering mechanisms,[78] tin selenide (SnSe) shows significantly different thermal behavior. Experimental studies confirm an ultralow κ_L of 0.23 ± 0.0 $3\mathrm{Wm^{-1}K^{-1}}$ in single crystal SnSe, attributing this low value mainly to strong lattice anharmonicity and structural anisotropy.[82]

To better understand this behavior, Carrete et al. used first principles DFT+BTE to perform detailed calculations of phonon dispersion relations, anharmonic IFCs, and phonon scattering rates. The results indicated a significant anisotropy in the phonon group velocities along the different crystallographic directions, which accounted for the observed directional dependence of the thermal conductivity of SnSe.[83] Third-order IFCs were calculated to investigate the impact of anharmonicity, which is crucial for evaluating the phonon-phonon scattering rates. Strong anharmonic interactions lead to high phonon scattering rates, effectively reducing phonon lifetimes. The strong anharmonicity is a factor for the exceptionally low κ_L of SnSe, thus preventing effective heat transfer across the lattice. This low κ_L of SnSe idea, which is dominated by strong intrinsic anharmonicity, makes it an ideal candidate for application in thermoelectrics. The suppressed thermal transport, coupled with its favorable electronic properties, results in an extremely high thermoelectric figure of merit (ZT), $\sim$ 2.6 at 923 K.[82] The dimensionless thermoelectric figure of merit is given by $ZT = S^2\sigma T/\kappa$, where S is the Seebeck coefficient, σ is the electrical conductivity, T is the absolute temperature, and κ is the total thermal conductivity comprising of both the lattice and electronic contribution. $ZT = 1$ is indicative of a more efficient thermoelectric material.[84,85]

Although studies of SnSe highlight the effects of strong anharmonic phonon interactions and anisotropy in limiting κ_L, however, studies from 2016–2025 have shown that the intricacy of heat transport in certain materials cannot be fully captured by second-order

perturbative treatments and traditional phonon-gas models.[79] This is particularly true in dynamically unstable or highly anharmonic crystals, where off-diagonal heat flux contributions and higher-order scattering must be considered.[86–88] The Wang et al. study of cesium chloride (CsCl) shows that the anharmonic effects and thermal resistance of CsCl are underestimated by the conventional three-phonon scattering approximations.[89] Employing an advanced BTE solver and first principles DFT, they incorporated four phonon scattering and self-consistent phonon (SCPH) renormalization, as well as the off-diagonal elements of the heat-flux operator, whose contributions are often overlooked in standard BTE implementations.[89]

Moreover, studies from 2016–2025 also demonstrate that the κ_L of many materials cannot be solely understood by the three-phonon process. Therefore, higher order phonon scattering, such as the four-phonon process, must be considered for some materials. Mathematically, the four-phonon scattering rate $\Gamma^{(4)}$ for a phonon mode λ, under Fermi's golden rule is expressed as follows:[90]

$$\Gamma_{\lambda}^{(4)} = \frac{2\pi}{\hbar^2} \sum_{\lambda'\lambda''\lambda'''} \left| V^{(4)}_{\lambda\lambda'\lambda''\lambda'''} \right|^2 \delta\left(\omega_{\lambda} \pm \omega_{\lambda'} \pm \omega_{\lambda''} \pm \omega_{\lambda'''}\right) \tag{9}$$

where $V^{(4)}$ is the fourth-order anharmonic interaction strengths and delta function is the energy conservation across all participating modes. Figure 3 shows a schematic of the four-phonon scattering processes. With the off-diagonal contribution, κ_L can be expressed as follows:[91]

$$\kappa_L{}^{\alpha\beta} = \kappa_d^{\alpha\beta} + \kappa_{nd}^{\alpha\beta} \tag{10}$$

where κ_d is the mode resolved diagonal phonon contribution, including both three-phonon and possible four-phonon interactions, and κ_{nd} is the coherence or mode coupling, the off-diagonal contribution.

2.2 Molecular Dynamics Simulation

The DFT+BTE approach, although accurate, is limited by the high computational expense, which is significantly involved in running complex calculations in evaluating atomic forces across numerous supercell configurations, substantially limiting its use for high-throughput

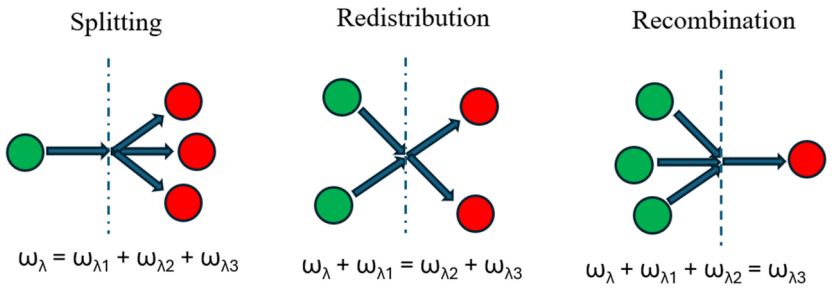

FIG. 3: Schematic of four-phonon scattering process (Reprinted with permission from the Royal Society of Chemistry, Copyright 2025)[1]

screening of new materials. In contrast, classical MD simulations provide a less expensive computational path for evaluating thermal transport properties.[92–95] In MD, atoms are modeled as interacting particles governed by Newtonian mechanics, where the energy and momentum are tracked over time by solving Newton's equation of motion.[1,96] The motion of each atom is determined by integrating Newton's second law, expressed by

$$m_i \frac{d^2 r_i}{dt^2} = F_i \tag{11}$$

where m_i is the mass of atom i, r_i is the position vector, and F_i is the net force acting on the atom i due to interactions with other atoms. Integrating Eq. (11) gives the velocity and position as a function of time, typically done using the finite-difference integration, such as Verlet algorithm. Thus, we can obtain the time evolution of atomic motion of the velocity and position at time $t + \Delta t$, where Δt is the timestep in an MD simulation, given the state of the system at time t.[97] In order to obtain the atomic trajectories, the force F_i must be determined, which results from the interatomic potential energy function $U(r)$. The potential $U(r)$, governs how atoms interact based on their positions, and the atomic force can be expressed by $F_i = -(\partial U/\partial r_i)$.

How the atoms attract and repel is determined by $U(r)$ and is central to MD simulations. The total potential energy (U) in classical MD is often decomposed into additive contributions from different orders of atomic interactions, as follows:[98]

$$\begin{aligned} U\left(1,\ldots,N\right) = &\sum_i U_1\left(r_i\right) + \sum_{i<j} U_2\left(r_i, r_j\right) \\ &+ \sum_{i<j<k} U_3\left(r_i, r_j, r_k\right) + \cdots + U_N\left(1,\ldots,N\right) \end{aligned} \tag{12}$$

where U_1, U_2, and U_3 are the single-body terms, two-body potentials (pairwise distance), and three-body potentials that capture the angular or directional interactions.

Physically meaningful force fields ranging from simple pairwise interactions, such as the Lennard-Jones potential, which can be employed in inert gas and colloidal systems,[99,100] and the Morse potential for covalent systems,[101,102] to more complex many-body potentials, such as the Tersoff[103] and Stillinger-Weber potential[98] for covalent systems, can be designed. Although in theory the potential energy expansion continues to four-body and higher order terms, in practice they are often rarely used primarily due to the high computational cost involved, as the number of interaction combinations increases with additional order. It is important to note that, for most materials, the inclusion of two- and three-body terms is sufficient in capturing the essential physics when the many-body effects are modeled implicitly within the potential form. However, MLP potentials like MACE[104] can effectively capture many-body interactions, bypassing the need to define high-order terms explicitly.

The technique utilized to extract physical properties is just as important as the selection and accuracy of the interatomic potential in MD simulations. Two primary methods are usually used for predicting thermal transport properties of inorganic crystals,

namely equilibrium molecular dynamics (EMD) and nonequilibrium molecular dynamics (NEMD), depending on whether the system stays in equilibrium with constant volume and energy or is driven out by external perturbation, such as heat flux or temperature gradient. Both methods vary in how transport properties such as thermal conductivity or diffusion are computed. EMD utilizes the Green–Kubo formalism to calculate the thermal conductivity, which relates it to the time integral of the heat current autocorrelation function at equilibrium. The κ_L in a GK-EMD simulation is expressed as follows:[105]

$$\kappa_L = \frac{V}{k_{\mathrm{B}} T^2} \int_0^{\infty} \langle J(0) \cdot J(t) \rangle \, dt \tag{13}$$

where V is the volume of the system, T is the absolute temperature, k_B is the Boltzmann's constant, $J(t)$ is the heat current vector at time t, and $\langle \cdot \rangle$ is the ensemble average.

In NEMD, thermal conductivity is determined by measuring the heat flux that results from applying a steady-state temperature gradient across the simulation cell. The thermal conductivity is determined using Fourier's law,[106] $\kappa = -(J/\nabla T)$, where J is the heat flux vector defined as the amount of energy transferred in a given time through a surface of a given area that is perpendicular to the direction of flux, and ∇T is the temperature gradient. In practice, a steady-state temperature profile is established by applying a heat source and a heat sink at opposite ends of the simulation domain, thus allowing the system to evolve until a stable thermal gradient is formed.

To perform MD simulations, a variety of software packages are commonly employed, including a large-scale atomic/molecular massively parallel simulator,[107] a Groningen machine for chemical simulation (GROMACS),[108] nanoscale molecular dynamics (NAMD),[109] DL_POLY,[110] HOOMD-blue,[111] and assisted model building with energy refinement (AMBER).[112] These packages rely on empirical or semi-empirical interatomic potentials, such as the Lenard-Jones, embedded atom method,[113] Tersoff,[103] or force fields like CHARMM[114] and AMBER.[112]

Salaway and Zhigilei[115] investigated heat transport in single-walled carbon nanotubes using the NEMD approach to calculate the thermal conductivity. Empirical Tersoff potential was utilized to model carbon-carbon interactions within the nanotubes with varying simulation parameters, such as lengths, cross-sectional area, and thermostat settings, thereby examining their influence on computed thermal conductivity using temperature gradients via Fourier's law. Their results revealed that shorter CNTs exhibited κ_L as a result of increased phonon-boundary scattering, while longer CNTs show higher thermal conductivities, reflecting the critical role of long wavelength phonons and the necessity of capturing phonon MFPs in low-dimensional systems. Several other studies have utilized EMD and NEMD in calculating the thermal conductivity of various materials,[105,116–121] thus showing the effectiveness of classical MD in capturing heat transport phenomena across different length scales. However, the accuracy is limited to the quality of the empirical interatomic potential used. To bridge the gap between the classical MD and first principles accuracy, more efforts have been channeled on the development of MLPs, which offer comparable DFT-level accuracy and significantly reduced computational cost (see more details in Section 3).

In contrast to classical MD, where predefined empirical potentials are usually used, *ab initio* molecular dynamics (AIMD) computes the interatomic forces on the fly based on quantum mechanical principles by DFT. This enables more accurate description of systems where electronic effects play an important role. The forces on the atoms are directly obtained from electronic structure calculations at each time step, thereby allowing accurate modeling of materials with complex or reactive bonding environments. Developments in AIMD since 2014 have been directed toward enhancing its computational efficiency and scalability without sacrificing its quantum-level accuracy. Notably, second-generation Car–Parrinello methods have decoupled the time evolution of electronic and ionic degrees of freedom, enabling longer simulation times and larger sytems.[122] Linear-scaling DFT and real-space representation such as CONQUEST and DFT-FE, have also enable AIMD to be applied to thousands of atoms.[123,124]

AIMD remains computationally expensive, especially for systems that require long time scales such as thermal transport or diffusion studies. Due to the high computational cost, DFT calculations are limited to short trajectories [$\sim$ (10–100 ps)] and modest system sizes [($< 10 \times 10^2$) atoms]. As a result of these constraints, AIMD is unsuitable for high-throughput screening or long-time dynamics and this has motivated the development of MLPs as efficient substitutes.[125]

3. AI/ML METHODS AS AN INNOVATIVE APPROACH FOR PHONON PROPERTIES

Predicting phonon properties of inorganic crystals, such as phonon dispersion, group velocity, Grüneisen parameters, and κ_L, has traditionally relied on first principles calculations, particularly DFT combined with anharmonic lattice dynamics or MD simulations. While accurate, these methods are generally computationally expensive, especially for large or complex systems. Since 2016, AI/ML have emerged as viable alternatives and accelerators for both DFT calculations and MD simulations. AI/ML techniques can either directly predict interested phonon properties or indirectly infer them.

The quality and scope of training data have a significant impact on the performance of AL/ML models in the prediction of phonon properties. Standardized, high-quality phonon datasets, which are essential for model development and benchmarking, can be obtained in several open databases, including phonondb@Kyoto University,[126] the Materials Project,[127] AFLOW,[128,129] OQMD,[130] and the NOMAD repository.[131] These sources offer phonon dispersion, DOS, and related thermal properties, which are often calculated using DFT. Aside from data availability, training strategies like multi-fidelity learning and transfer learning are very frequently used to maximize predictive accuracy with minimal high-fidelity data. Typically, the workflow involves pretraining on large low-fidelity datasets (e.g., coarser DFT or empirical potential calculations) and fine-tuning on smaller, high-fidelity datasets (e.g., highly converged DFT or experimental results).[132–136] Using this approach will thus enable models to capture a wide range of structure-property trends, while also adapting to the precision of costly calculations, therefore making it particularly valuable for high-throughput phonon transport screening in data-scarce regimes.

Section 3 is divided into two parts: direct AI/ML methods, in which AI/ML models are trained to learn phonon-related quantities directly from training data, i.e., the famous atomic structure-property relationship, and indirect AI/ML methods, where bottom information are predicted first, such as atomic forces in displaced supercells, and then the traditional workflow of the DFT+BTE approach can be used to predict all phonon properties, such as group velocities, mean square displacement (MSD), three phonon scattering phase space (P_3 parameter), κ_L, etc.

3.1 Direct AI/ML Algorithms and Methods

In direct methods, AI/ML algorithms are trained to predict specific phonon-related quantities from atomic or crystallographic input data. These approaches aim to bypass the full DFT workflow by directly learning the complex relationship between atomic structure and phonon transport behavior. Traditional ML models and GNNs have been used extensively over the years to directly predict the phonon properties, especially the κ_L of inorganic crystals with high accuracy. The traditional AI/ML models require input representation called descriptors or features from which the models learn the relationship with the corresponding target property. The accuracy of these traditional ML models largely depends on the effective input representation of the crystal structures.[137] Selecting good or effective descriptors is an important step in directly training the ML model to well predict a target property. The following are some of the characteristics when selecting a descriptor:[37]

1. Low Dimensionality: The dimensionality of the descriptors should be as low as possible. Lower-dimensional descriptors reduce the amount of training data required, facilitating more efficient model training.

2. Uniqueness and Relevance: The descriptor should uniquely characterize both the crystalline materials and the property-relevant elementary processes. Ideally, there should be a one-to-one correspondence between the input (descriptor) and the output (material property), making the ML model easier to train and more accurate.

3. Discriminative Power: Materials that are significantly different (or similar) should have correspondingly different (or similar) descriptor values. This helps the ML model adjust its internal parameters effectively to recognize underlying patterns and relationships.

4. Computational Efficiency: The calculation of the descriptor should be significantly less computationally intensive than that required for evaluating the target property, since the trained ML models will usually be used for subsequent prediction and screening of large-scale unknown materials. Simple calculation or evaluation of corresponding material descriptors will facilitate such a process. This is especially important for phonon property prediction, as traditional approaches, such as DFT+BTE, are highly time- and resource-consuming.

The Pearson correlation coefficient, ranging from –1 to +1, is a commonly used metric for assessing the strength and direction of a linear relationship between two variables. A value of –1 indicates a strong negative correlation, whereas +1 indicates a strong positive correlation. Ojih et al.[138] have shown Pearson correlation for some typical material descriptors with κ_L, as shown in Fig. 4. Several descriptors exhibit considerably negative correlations with κ_L, such as the volume of the primitive cell (–0.49), bond length (–0.45), total atomic weight (–0.21), and average number of electrons (–0.18). These values suggest that higher values of these descriptors are associated with lower κ_L. This trend is physically reasonable, as larger cell volumes or longer bond lengths typically imply weaker interatomic bonding, which leads to lower phonon group velocities and thus lower κ_L. Conversely, atom number density, the number of unpaired electrons, and mass density

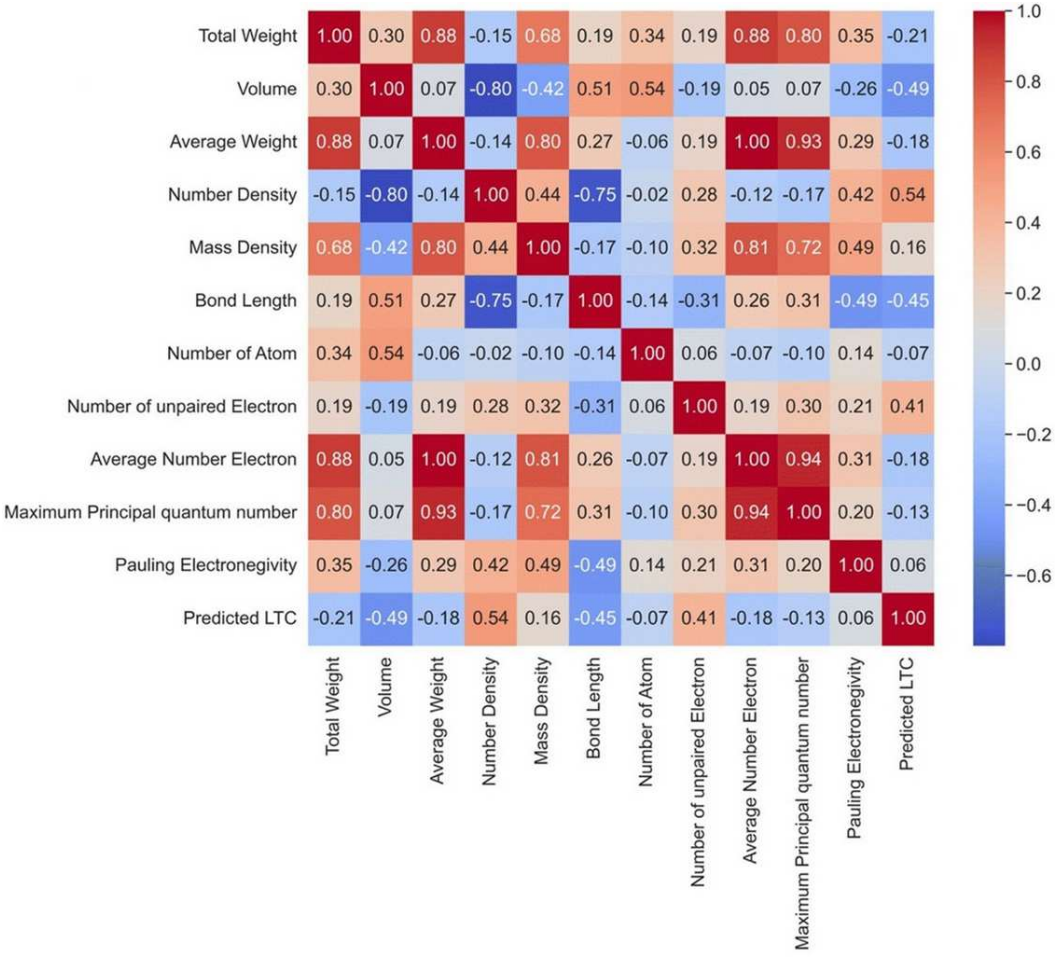

FIG. 4: Pearson correlation between material descriptors and κ_L of inorganic crystals (Reprinted under a Creative Commons Attribution–NonCommercial 3.0 Unported License, Copyright 2024; http://creativecommons.org/licenses/by-nc/3.0/)[138]

show positive correlations with κ_L, with Pearson correlation values of 0.54, 0.41, and 0.16, respectively. A higher atom number density generally reflects tighter atomic packing, which corresponds to stronger interatomic bonds and higher phonon group velocities, thereby enhancing phonon-mediated energy transport in the lattice.

In contrast, descriptors such as the total number of atoms, the maximum principal quantum number, and Pauling electronegativity exhibit negligible or weak correlations with κ_L. The identification of both positively and negatively correlated descriptors is expected to facilitate the accelerated screening of large-scale hypothetical structures in future materials discovery efforts.

Carrete et al.[41] and Liu et al.[139] used traditional ML methods to directly predict the κ_L of half-Heusler structures by training on DFT κ_L data. Qin et al.[140] trained 15 ML models from fundamental material descriptors to predict the κ_L of materials, where 80% of the data were used for training and the remaining 20% for testing. The study demonstrated that the traditional ML models could effectively capture the correlation between material properties and κ_L, enabling accurate predictions across diverse materials. Table 1 shows a comparison of evaluation metrics for all 15 models. It was found that the long short-term memory network has the highest performance in training and prediction in terms of lowest root mean squared error (RMSE) and mean absolute error (MAE) of testing data, highest R^2 score. However, this algorithm has the lowest computing speed (several folders to even one order of magnitude longer than other ML models, as seen in Table 1).

TABLE 1: Comparison of evaluation metrics for all 15 ML models for predicting κ_L (Reprinted with permission from the Royal Society of Chemistry, Copyright 2023)[140]

ML models	RMSE of test set	R^2 of test set	MAE of test set	Time cost(s)
Linear	26.493	0.8096	13.5803	6.49
Ridge	26.3697	0.8103	13.5384	3.46
SGD	17.4929	0.8241	9.5713	3.58
LinearSVR	11.9823	0.8479	6.8762	3.68
SigmoidSVR	20.6119	0.3432	9.7734	1.94
rbfSVR	14.6274	0.7547	6.9582	1.80
PloySVR	12.0221	0.7496	7.2365	3.61
Decision tree	19.378	0.5348	8.6358	4.56
DGBT	10.3623	0.8158	6.957	5.69
RF	9.6385	0.8767	6.0574	3.76
LightGBM	12.9994	0.7398	7.7365	4.56
ANN	8.7211	0.8593	5.7933	18.19
CNN	8.4061	0.8799	5.1674	19.57
RNN	8.3726	0.8748	5.3209	61.03
LSTM	8.3593	0.8866	5.4011	125.46

Chen et al.[141] proposed a data-driven framework for predicting the κ_L of inorganic materials using traditional ML techniques. They compiled a dataset of $\sim$ 100 inorganic solids with experimentally measured κ_L values and employed various regression models, with a particular focus on Gaussian process regression (GPR). The descriptors used to describe the materials included fundamental physical and chemical properties such as atomic mass, bond length, unit cell volume, and Debye temperature, which were extracted from either first principles calculations or experimental databases. Feature selection and dimensionality reduction techniques, such as principal component analysis (PCA), were applied to identify the most relevant descriptors for κ_L prediction. The GPR model achieved strong predictive performance, with a low MAE and high correlation coefficient on the test set, outperforming traditional physics-based models, such as the Slack formula. The study also conducted uncertainty quantification to evaluate model robustness. Importantly, the authors demonstrated the model's ability for generalization by applying it to screen a large database of hypothetical compounds, identifying potential materials with low or high κ_L. Despite the small dataset used, this work highlights the potential of traditional ML approaches in accelerating thermal materials discovery and offers a practical workflow that combines domain knowledge, data curation, and statistical modeling for κ_L prediction.

Tranas et al.[142] focused on predicting the κ_L of half-Heusler compounds using a combination of DFT, random forest (RF) regression, PCA, and active learning. The goal is to enhance model accuracy while reducing the number of costly DFT calculations. The workflow begins by calculating a small number of κ_L values using DFT for a subset of half-Heusler compounds. These values, along with descriptors, such as atomic radius, electronegativity differences, and crystal structural features, serve as the training data for a RF regression model. PCA was applied to reduce descriptor dimensionality and capture the most significant variance in the data. To further improve efficiency and model generalization, the authors incorporated active sampling, a form of active learning. Here, instead of selecting training samples randomly, they iteratively chose new data points based on the model's uncertainty and PCA space coverage. This was to ensure a more diverse and informative training set, leading to better model performance with fewer training samples. The study shows that with active learning, the RF model can achieve accurate κ_L predictions across a large chemical space while minimizing the number of DFT calculations needed. The final model can rapidly screen half-Heusler compounds and identify promising candidates for thermoelectric applications.

The GNN models, on the other hand, combine the descriptors and learning model into one inseparable step, i.e., the model learns material properties directly from the connection of atoms in the crystal. The crystal graph convolutional neural network (CGCNN) developed by Xie and Grossman was one of the first GNN models to represent periodic materials using graph structures.[143] Although initially developed for electronic properties, later adaptations have shown that it can be used to predict phonon DOS and related thermal properties. Other GNN models have been developed over the years with each having unique descriptors from the crystal structures to improve prediction. Some representative GNN models in this line include atomistic line GNN (ALIGNN),[144] orbital graph

convolution neural network (OGCNN),[145] global attention GNN (deeperGATGNN),[146] Improved CGCNN (iCGCNN),[147] Materials Graph Network (MEGNet),[148] etc.

Ojih et al.[138] trained different types of GNN models for predicting κ_L with high accuracy. Figure 5 shows the training and testing data for their work. While the deeperGATGNN model has the highest performance for training data in terms of lowest MAE and highest R^2 score, the performance for testing data do not show very high performance. In contrast, the ALIGNN model has excellent overall performance for both training and testing data, and therefore this model is recommended for further development and training if more data are available. This work also suggests that more research effort is needed for further evaluation of various GNN models on an even larger scale of training data in order to cover diverse material compositions and symmetries, due to the highly nonlinear relationship between atomic structures and thermal transport properties.

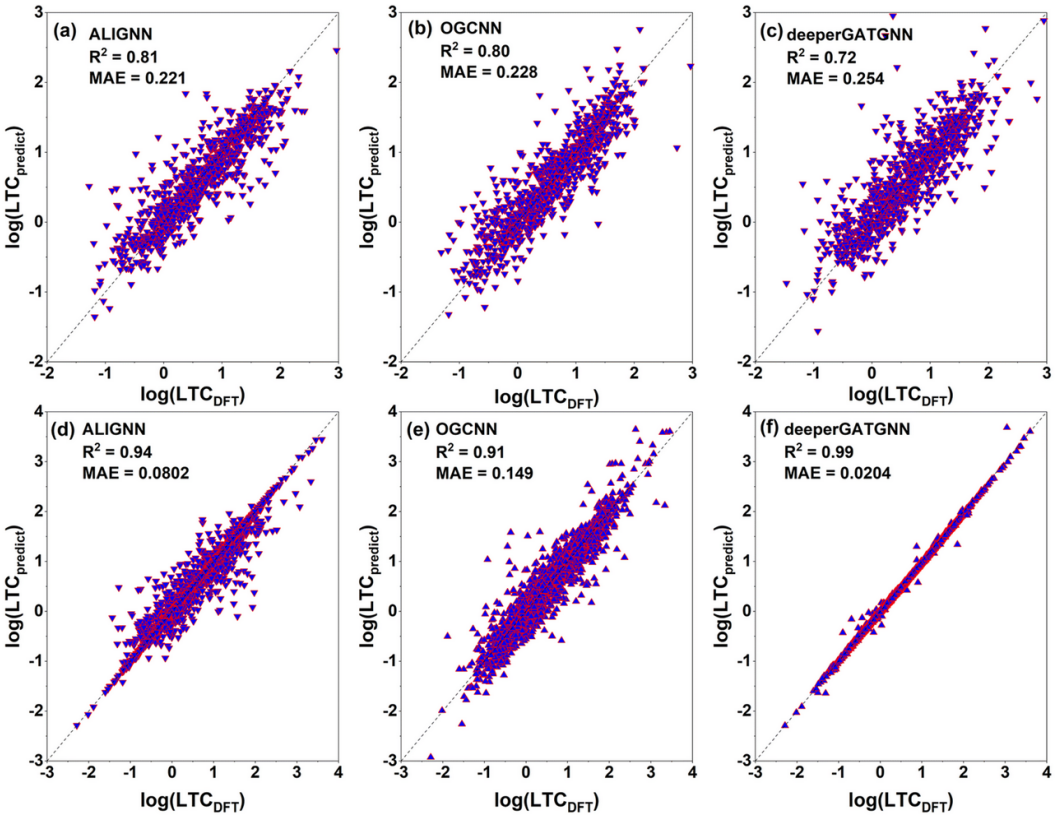

FIG. 5: Testing (top panels) and training (bottom panels) results of κ_L for the three predictive GNN models for 942 and 3769 structures, respectively, are as follows: (a) ALIGNN, (b) OGCNN, (c) deeperGATGNN, (d) ALIGNN, (e) OGCNN, and (f) deeperGATGNN. The ALIGNN model shows the best performance for the testing (unseen) data (Reprinted under a Creative Commons Attribution–NonCommercial 3.0 Unported License, Copyright 2024; http://creativecommons.org/licenses/by-nc/3.0/).[138]

Ojih et al.[149] trained a deeperGATGNN[146] model on five mechanical properties, namely, bulk modulus, shear modulus, Young's modulus, Poisson ratio, and hardness using 10,158 elastic constants by DFT calculations as training data. It is known that the κ_L of a crystalline material is proportional to the phonon group velocity, and the phonon group velocity is proportional to Young's modulus and bulk modulus, shown as follows:

$$k = \sum C_i v_i \tau_i, \quad v \propto \sqrt{\frac{E}{\rho}}, \quad B = \frac{E}{3\,(1-2v)} \tag{14}$$

where ρ, B, E, and v are the mass density, bulk modulus, Young's modulus, and Poisson's ratio, respectively. This gives vital information that κ_L is proportional to bulk modulus. Ojih et al.[149] used this information and predicted the bulk modulus of 775,947 inorganic crystal structures from the open quantum material database (OQMD) using ALIGNN and verify the κ_L of 338 materials randomly selected from the predicted bulk modulus, as shown in Fig. 6. It can be clearly seen that the κ_L is proportional to the bulk modulus. It is worth pointing out that calculating bulk modulus of a crystal by direct DFT calculations through elastic constants is not very heavy, which suggests that the bulk modulus can be used as a descriptor for screening extremely high or low κ_L materials through high-throughput DFT calculations.

Ojih et al.[150] also used the MSD and P_3 parameters to investigate the κ_L of materials. Both the MSD and P_3 parameters were trained on 4041 DFT-κ_L data and used to predict on 22,899 dynamically stable structures to get structures with low κ_L. Validation was done on

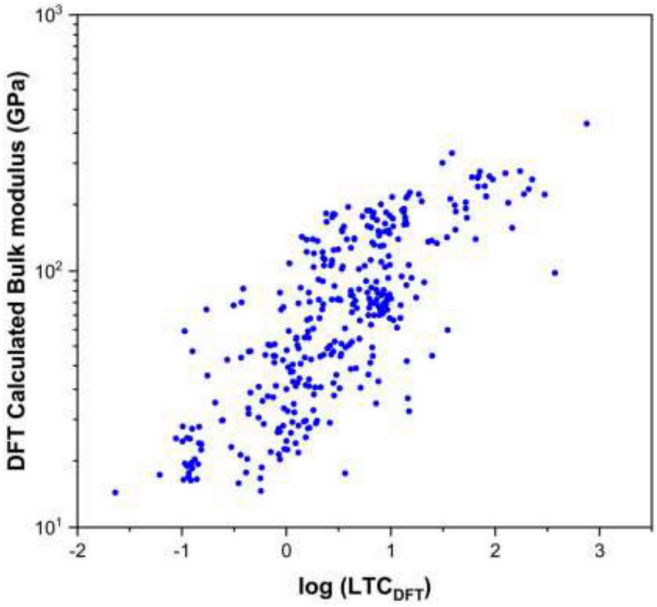

FIG. 6: Correlation between bulk modulus and κ_L for 338 recommended structures (Reprinted under a Creative Commons Attribution 4.0 International License, Copyright 2023; https://creativecommons.org/licenses/by/4.0/)[149]

359 structures, as shown in Fig. 7. It is found that κ_L inversely correlates with both MSD and P_3 parameters, meaning high MSD and high P_3 parameters usually correspond to low κ_L, and vice versa. Such correlation is consistent with phonon transport physics considering that high MSD corresponds to soft lattices and usually strong phonon anharmonicity, while a high P_3 parameter corresponds to a large space of three-phonon scattering, both of which therefore lead to low κ_L. This study demonstrates that the large P_3 parameter and large MSD are good material descriptors for quick screening materials with ultralow κ_L. It is worth pointing out that both the P_3 parameter and MSD can be calculated by the second-order IFCs and well trained by medium-size DFT data. Therefore, these two material descriptors provide an efficient approach for future ML model development for phonon transport property.

While predictive accuracy is a top priority, the interpretability of AI/ML models is equally important for advancing the physical understanding of phonon transport. For descriptor-based models described herein, where features like primitive cell volume, bond length, atomic mass, or MSD are explicitly defined, interpretability is accessible via correlation analysis, feature importance rankings from tree-based algorithms, or Shapley additive explanation values.[148,151,152] Not only do such analyses confirm established physical relationships (e.g., the inverse correlation between κ_L and cell volume), but they can also potentially uncover unexpected trends that may seed new hypotheses regarding structure–property relationships. For end-to-end methods, such as the GNN models described here, interpretability methods involve attention weight analysis, node/edge saliency mapping,

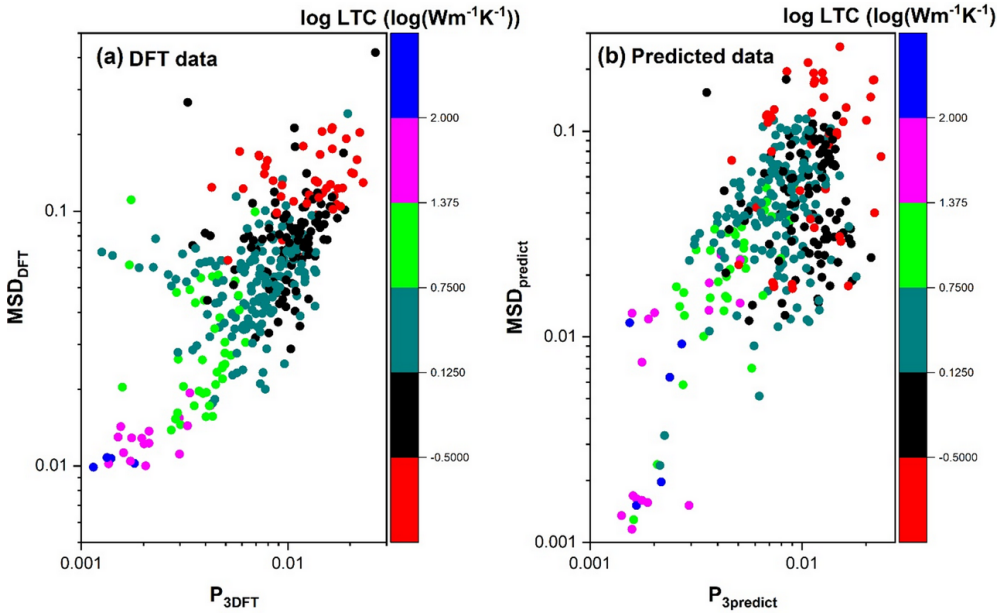

FIG. 7: Correlation between bulk modulus and κ_L for 338 recommended structures (Reprinted under a Creative Commons Attribution 4.0 International License, Copyright 2023; https://creativecommons.org/licenses/by/4.0/)[149]

and latent-space visualization.[143,153] These analyses have the potential to indicate which atomic environments, bonding motifs, or symmetry features most strongly determine the models' prediction of phonon-related properties. In several instances, GNN-learned embeddings have been correlated with physical quantities like bond stiffness or anharmonicity,[154] thus bridging the gap between abstract learned representations and concrete phonon physics. Including such interpretability analyses not only enhances confidence in AL/ML predictions but also offers a complementary path toward uncovering and refining the underlying mechanisms of thermal transport.

3.2 Indirect AI/ML Algorithms and Methods

MLPs are a key subclass of the "indirect" AI/ML approach, as they replace expensive *ab initio* force evaluations with fast surrogate models while retaining compatibility with established lattice dynamics and BTE workflows. Widely used MLP families include the Behler–Parrinello high-dimensional neural network potential,[155] the Gaussian approximation potential with smooth overlap of atomic positions (SOAP) descriptors,[156,157] the moment tensor potential (MTP),[158,159] the spectral neighbor analysis potential,[160] and the deep-learning potential package DeepMD/DeePMD-kit.[161,162] From 2018 onwards, E(3)-equivariant GNN models, such as NequIP[163] and the atomic cluster expansion (ACE) and its modern variants (MACE),[104,164] have been developed to improve data efficiency and transferability.

Several comprehensive reviews summarize the theoretical foundations, training strategies, and applications of MLPs to materials simulations.[165–167] In the context of phonon and thermal transport, MLPs have been successfully applied in combination with Sheng-BTE for high-throughput lattice thermal conductivity (κ_L) prediction[168] and in Green–Kubo molecular dynamics for strongly anharmonic systems such as zirconia,[169] achieving near-DFT accuracy with orders-of-magnitude reductions in computational cost. In general, the indirect approach does not predict phonon properties directly from material descriptors. Instead, it focuses on learning intermediate physical quantities—most commonly interatomic forces (bottom information) or upper-level parameters such as IFCs. These quantities are then fed into established physics-based frameworks (e.g., anharmonic lattice dynamics plus BTE) to derive κ_L and other phonon-related properties. This AI/ML method effectively replaces the most computationally expensive step in the DFT+BTE workflow (i.e., force calculations as indicated by the center vertical arrow in Fig. 8), while maintaining the rigor and interpretability of traditional DFT+BTE method.

A simpler version for replacing atomic force evaluation by ML model in the traditional DFT+BTE workflow or MD simulations can be easily realized for a single material or some limited number of materials or material families. Such a method is usually called an MLP or machine learning force fields (MLFFs). There are many studies in this area since the high-cost DFT calculations are usually much smaller than the abovementioned universal MLPs that try to cover very broad material compositions and symmetries. Korotaev et al.[170] developed an MLP and trained it on DFT data to model the lattice dynamics of complex compounds, specifically skutterudites. By employing active learning strategies,

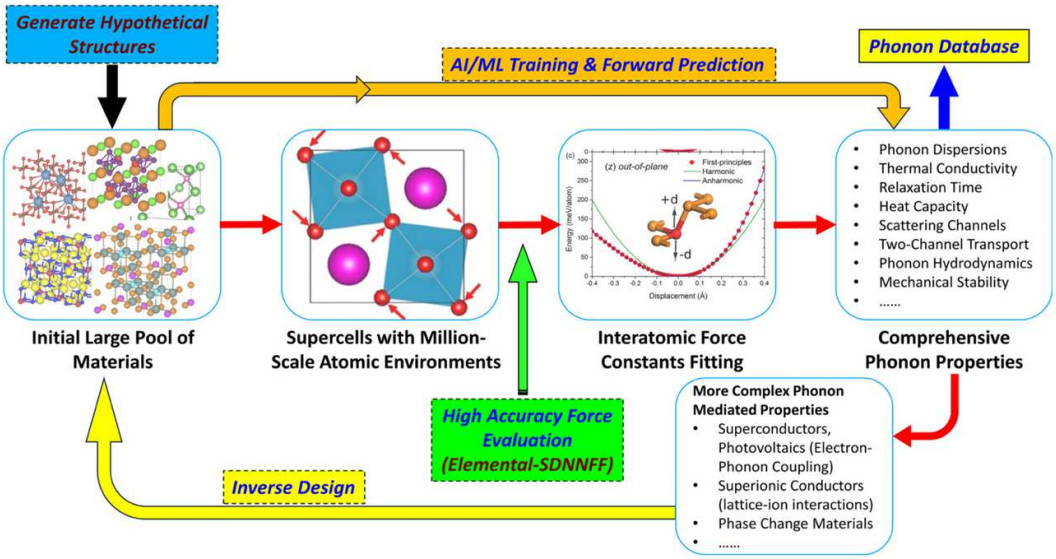

FIG. 8: Schematic of the deployment of indirect AI/ML algorithms and methodologies (center vertical arrow) in phonon transport. The middle horizontal arrows indicate the traditional DFT+BTE approach for phonon property prediction. The top horizontal arrows denote the direct AI/ML model for structure-phonon property relationship and forward prediction of phonon transport properties (Reprinted under a Creative Commons Attribution 4.0 International License, Copyright 2024; https://creativecommons.org/licenses/by/4.0/).[37]

they minimized the number of expensive DFT calculations required for training. The MLP accurately reproduced vibrational spectra and κ_L values, demonstrating that accurate and reliable potentials can be obtained with a limited number of quantum-mechanical calculations. This approach significantly reduces computational costs while maintaining high accuracy in predicting thermal properties.

Srivastava et al.[171] introduced a ML-assisted method to extract anharmonic force constants efficiently. Applying their approach to a dataset of 220 ternary materials, they achieved a reduction in computational time from 48×10^4 CPU hours to $< 12 \times 10^3$ CPU hours, while maintaining κ_L prediction accuracy within 10%. This significant acceleration facilitates high-throughput screening of materials for thermal applications. Lu et al.[172] combined on-the-fly MLFFs with MD simulations to study the κ_L of $ZrSe_2$. This hybrid approach allowed them to capture anharmonic phonon interactions effectively, leading to accurate κ_L predictions. The methodology enables the analysis of large supercells and long-wavelength phonons, which are challenging for traditional *ab initio* methods. Ouyang et al.[173] developed an MLP based on a matrix tensor algorithm to study phonon anharmonicity in materials like cubic BAs and diamond. Their approach accurately captured phonon mode softening and linewidth broadening induced by anharmonicity at finite temperatures. The ML potential enabled efficient MD simulations, providing insights into thermal transport properties with reduced computational effort.

In the indirect ML approach to predict phonon properties of large-scale materials instead of a single material or family, the process begins by generating a dataset of atomic structures where accurate atomic force information is needed (Fig. 8). These structures may come from random displacements around equilibrium positions, snapshots from MD simulations, or known crystalline configurations. For a selected subset of these structures, DFT calculations are performed to compute the atomic forces, which serve as high-fidelity reference data. Each atom's local environment is converted into a set of descriptors that capture its chemical and structural context. These descriptors may be handcrafted, such as symmetry functions or SOAP, or they may be learned automatically using ML architectures like GNNs. The ML model is then trained to learn the mapping from atomic environments to the corresponding atomic force vectors. The model is optimized to minimize the prediction error when compared to the DFT-calculated forces.

Once trained, the ML model is used to predict atomic forces for new, unseen structures, or displaced configurations of a target material. Using these ML-predicted forces, one can reconstruct the IFCs using FDMs or regression-based fitting schemes. These IFCs capture the harmonic and anharmonic interactions between atoms and are essential inputs for subsequent calculations of full phonon properties. The reconstructed IFCs are then fed into lattice dynamics tools, such as Phonopy,[58] which computes phonon dispersion relations, group velocities, and heat capacities, and into thermal transport solvers like ShengBTE,[62] which calculates phonon scattering rates and ultimately solve phonon BTE. From this, the κ_L is obtained. This indirect workflow allows researchers to retain the accuracy of traditional DFT-based approaches while significantly reducing computational costs by replacing expensive force calculations with fast, ML-based predictions.

Beyond accelerating the computation of IFCs, ML can also be employed at later stages of the phonon transport workflow, particularly for expediting the evaluation of phonon scattering rates. This is a critical step in solving the phonon BTE and can become a computational bottleneck, especially when higher-order processes are included. Since 2021, studies have shown that ML can dramatically accelerate the calculation of phonon scattering rates, a critical bottleneck in predicting κ_L, particularly when considering higher-order processes such as four-phonon scattering. Guo et al. proposed a surrogate ML model capable of predicting phonon scattering rates and κ_L with high accuracy compared to first principles direct approaches, up to two orders of magnitude reduction in computational cost.[174] In a related work, Guo et al. proposed a maximum-likelihood estimation strategy to efficiently sample phonon scattering events, thus enabling highly converged thermal conductivity calculations using minimal computational resources.[175] These advances are indicative of a promising path for high-throughput predictions of thermal transport in complex materials.

Various ML models, such as CHGnet,[176] MACE,[104] etc., can be used for predicting atomic forces in displaced supercells. Rodriguez et al.[37,177,178] developed an ML model called the Elemental Spatial Density Neural Network Force Field (Elemental-SDNNFF) to predict atomic forces with near-DFT accuracy. Elemental-SDNNFF distinguishes itself through a novel 3D mesh of density functions that collectively map the atomic environment, offering a physically intuitive representation of forces exerted on the central atom.

The efficiency and precision of the Elemental-SDNNFF approach arise from three key innovations: (i) eliminating reliance on the chain rule for total energy derivatives, a major source of error in ML-derived atomic forces; (ii) training the model using only a central atom and its local environment, enabling a scalable $n \times m$ growth of training data, where n is the number of atoms in a supercell and m is the number of structures evaluated via first principles methods; and (iii) drastically reducing both the parameter count and manual effort required to train a neural network model that achieves convergence in per-atom properties. Trained on 9.4 million atomic environments using active learning, the model was applied to 11,866 Heusler structures,[178] including full, half, and quaternary types. The predicted atomic forces were used in the traditional DFT-based phonon workflow to calculate phonon properties, such as κ_L. The study identified 774 Heusler structures with ultralow κ_L (< 1 $\mathrm{Wm^{-1}K^{-1}}$). This work showcases a scalable, indirect ML approach for high-throughput materials discovery and also highlights how p–d orbital hybridization plays a key role in reducing κ_L.

Rodriguez et al. further demonstrate the Elemental-SDNNFF approach to predict atomic forces across $\sim (8 \times 10^4)$ cubic crystals spanning 63 elements (Fig. 9).[179] A key strength of the Elemental-SDNNFF framework is its ability to resolve phonon transport physics with parity to first principles accuracy, enabling simultaneous prediction of diverse phonon properties through a unified model. Trained on 3182 first principles datasets and applied to 77,091 unexplored structures, the method identifies 13,461 dynamically stable cubic systems with ultralow κ_L (< 1 $\mathrm{Wm^{-1}K^{-1}}$), including 36 structures validated against first principles benchmarks. On the basis of the screening of such large amount of ultralow κ_L materials, they further propose two computationally efficient descriptors—mean square displacement and bonding-antibonding metrics—to streamline the identification of ultralow κ_L materials, reducing reliance on costly DFT simulations. The model also quantifies the interplay between off-diagonal coherence and diagonal phonon populations, capturing the transition from particle-like to wave-like heat transport. The Elemental-SDNNFF framework holds significant promise for accelerating the discovery of advanced phononic materials for applications in thermoelectrics, superconductivity, and topological phonons for quantum technologies.

4. CHALLENGES AND OPPORTUNITIES FOR AI/ML-DRIVEN RESEARCH FOR PHONONS

Although data-driven AI/ML methods offer promising avenues for reducing the computational cost of *ab initio* phonon calculations (e.g., DFT+BTE) and predicting phonon properties, several challenges hinder their widespread adoption.

4.1 Data Limitations: Quality, Quantity, and Accessibility

Phonon properties are inherently complex, with highly nonlinear and implicit relationships to material structure. Generating sufficient high-quality training data to cover diverse material families and symmetries remains a bottleneck. Although DFT+BTE remains the gold

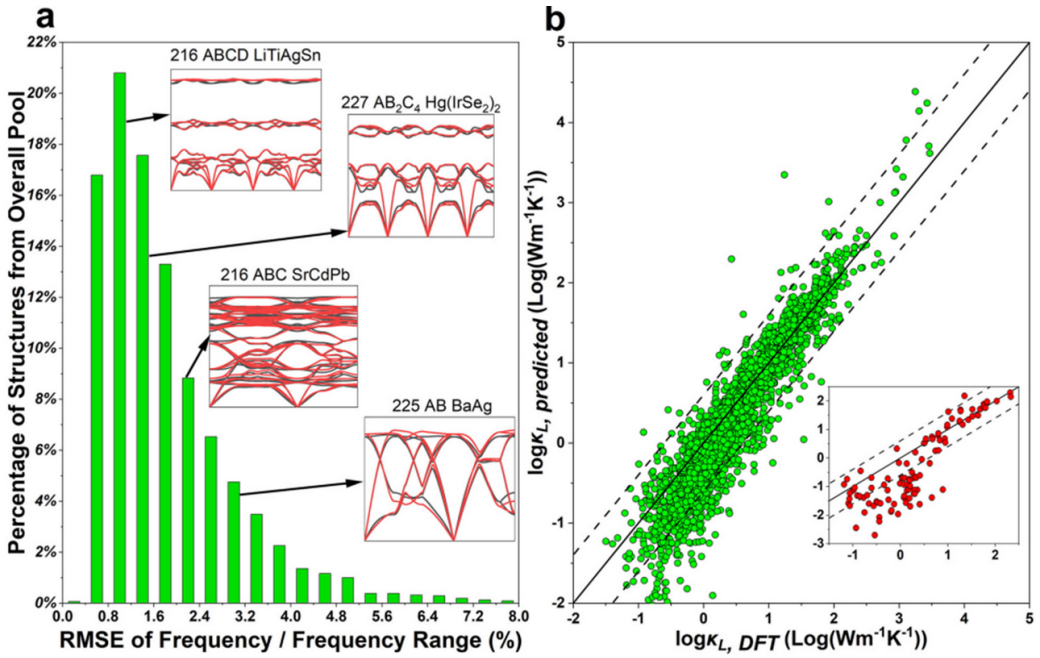

FIG. 9: (a) Comparison of the RMSE of phonon frequency normalized by the structure's specific frequency range. (Insets) Phonon dispersions linked to the relative error containing DFT and prediction for visualization. (b) Lattice thermal conductivity at 300 K between DFT and the developed single neural network Elemental-SDNNFF model for 3107 stable structures predicted by DFT. (Inset) The comparison between the predicted and DFT LTC of 64 untrained structures on the same scale (Reprinted under a Creative Commons Attribution 4.0 International License, Copyright 2023; http://creativecommons.org/licenses/by/4.0/).[179]

standard for data accuracy, progress hinges on building robust phonon databases. A centralized, publicly accessible phonon database (Fig. 8) could revolutionize the field by: (i) providing standardized, curated phonon data (band structures, lifetimes, scattering rates) alongside existing material databases, such as Materials Project, ICSD, OQMD; (ii) offering interactive tools for data visualization, comparison, and user-contributed DFT results; and (iii) facilitating global collaboration via user-friendly upload protocols and feedback channels, accelerating data accumulation, and enabling cross-disciplinary innovation.

4.2 High Computational Cost MD

From a computational cost perspective, machine learning interatomic potentials (MLPs) offer orders-of-magnitude speedup over AMID by replacing on-the-fly quantum mechanical force calculations with pretrained surrogate models, enabling simulations of much larger systems and longer time scales. However, generating high-quality training datasets—often requiring extensive AIMD or DFT calculations—can itself be computationally

expensive, particularly when diverse chemical environments, defects, or wide thermodynamic ranges must be represented. Moreover, in terms of large-scale MD simulations equipped with MLPs, additional cost-related limitations arise, including substantial memory and storage demands for millions of atoms and long trajectories, communication overhead and load imbalance in massively parallel computing, the need for femtosecond timestep to maintain numerical stability, and the computational burden of post-processing massive datasets. Also, finite-size effects may still necessitate further scaling to capture long-wavelength phonon contributions, which can offset some of the efficiency gains offered by MLPs.

4.3 Constraints in Modeling Thermal Transport in Complex Crystal Systems

Another important limitation lies in the intrinsic complexity and diversity of crystal systems, which poses a fundamental challenge for direct ML predictions of thermal conductivity. Thermal transport in solids arises from a highly material-specific interplay of factors, including chemical bonding characteristics, lattice symmetry, phonon dispersion relations, higher-order anharmonic processes such as four-phonon scattering, and even contributions from non-propagating vibrational modes (diffusions). These mechanisms can differ substantially between materials, meaning that patterns learned from one subset of compounds may not transfer reliably to others with different bonding types or dominant heat-carrying processes. Consequently, the confidence in applying direct ML models—particularly those trained on limited or compositionally biased datasets—to predict the thermal conductivity of entirely new materials remains uncertain. Future work should aim to quantify and improve this generalization capability, for example, by expanding training data diversity, incorporating physically informed descriptors, and systematically evaluating model performance across contrasting thermal transport regimes.

4.4 Bridging Multi-Resolution Data Gaps

Current AI/ML models for phonon properties rely on limited datasets (often just hundreds of κ_L values) due to the high computational cost of *ab initio* methods. Conversely, approximate theoretical models generate abundant but lower fidelity data. Hybrid approaches that integrate multi-resolution data (combining high-accuracy DFT and scalable theory-based datasets) could train ML models to balance speed and accuracy. Such models would better generalize to unexplored materials, addressing the critical "extrapolation problem," where ML struggles beyond training domains.

4.5 Developing and Comprehensively Evaluating Universal Machine Learning Potentials

Though MLPs accelerate phonon predictions by orders of magnitude, most are narrowly tailored to specific materials or families. Training universal MLPs (uMLPs) capable of handling diverse chemistries, symmetries, and atomic species requires massive datasets.

Recent efforts (2021–2023; e.g., CHGNet,[176] M3GNet,[180] MACE,[104] and EquiformerV2[181]) leverage large-scale DFT datasets (e.g., Materials Project, ICSD, and OQMD) to pretrain broad MLPs. First, we need to understand the accuracy of existing uMLPs for phonon transport property prediction. Especially related to phonons, the accuracy of atomic force evaluation by the uMLPs is the most concern, either for replacing force evaluation in the DFT+BTE approach or for running MD simulations. Efforts in this regard have just begun.[182] Second, to overcome the impossibility of exhaustive DFT sampling, active learning could identify critical gaps in training data, guiding targeted DFT calculations to refine uMLPs iteratively. With speeds nearing empirical models, uMLPs hold immense potential for high-throughput screening.

4.6 Enhancing Generalization and Inverse Design

Most AI/ML models excel at interpolation (predicting within known data regimes) but falter at extrapolation. Discovering novel materials with exceptional phonon properties demands generative models capable of inverse design—proposing structures with target phonon behaviors (Fig. 8). Emerging tools like Google DeepMind's GNoME, which claims to predict more than 38×10^4 stable materials via active learning and GNNs, highlight progress. However, questions remain about dynamic stability (e.g., phonon-mediated instabilities), underscoring the need for AI/ML frameworks that uncover hidden physical laws rather than relying on interpolation. For example, many ML-enabled crystal structure prediction models failed to generate dynamically stable materials with a high success rate, which could waste lots of CPU time on filtering out the unstable structures by subsequent DFT calculations. Therefore, implementing dynamic stability into the hypothetical structure generation from the very beginning is an urgent task to do in this field.

4.7 Unlocking Hidden Physics and Chemistry of Phonons from Data

AI/ML can also enable the discovery of hidden physics and chemistry rules and principles by analyzing complex, high-dimensional datasets to uncover patterns and relationships imperceptible to traditional methods. By training on experimental or simulated data, AI/ML models, such as neural networks or graph-based algorithms, can infer latent governing equations, material properties, or reaction mechanisms, even in systems with incomplete theoretical frameworks. Techniques such as symbolic regression,[183] unsupervised learning, or attention mechanisms help distill interpretable rules or features, revealing underlying principles of molecular interactions, phase transitions, or energy dynamics, all of which could be linked or related to phonon transport in inorganic crystals. For example, how can the AI/ML models help us with extracting physics and chemistry rules for dynamic stability, i.e., generating new materials free of negative vibrational frequencies in the Brillouin zone? Is it possible to get new rules for materials with strong phonon anharmonicity beyond those mechanisms that have already been understood such as lone-pair electrons, rattling effect? The data-driven approach is expected to accelerate hypothesis

generation, bridge gaps in multiscale modeling, and unlock predictive insights for novel materials, chemical processes, or physical phenomena.

4.8 Expanding Current AI/ML Approach Beyond Single Crystals and Phonons

Real-world applications (e.g., thermoelectrics) require understanding thermal transport in heterogeneous systems (interfaces, disordered materials). Current AI/ML models, trained on single-crystal phonon data, struggle with symmetry-breaking interfaces or aperiodic structures. Universal MLPs could enable large-scale MD (e.g., NEMD) to study interfacial thermal conductance (ITC). Hybrid workflows combining ML-predicted phonon DOS overlap and semi-empirical models (e.g., diffuse mismatch model in almaBTE) may accelerate interface screening. ML-driven studies of disordered superlattices have already challenged conventional wisdom, revealing optimized phonon suppression in aperiodic systems. AI/ML models for ITC prediction and screening based on single crystal DFT-IFC data have been developed and demonstrated on high electron mobility transistors with wide bandgap materials.[184,185]

As phonons play central roles in many societally important science and technology applications, it is then intuitive to expand the current AI/ML approach beyond phonon itself. Phonon-mediated superconductivity arises from the interaction between electrons and lattice vibrations, where phonons facilitate the attractive force that binds electrons into Cooper pairs, enabling resistance-free electrical current. This mechanism, central to conventional superconductors described by Bardeen-Cooper-Schrieffer theory, hinges on the strength of electron-phonon coupling and the material's phonon spectrum. Since 2015, advances in high-pressure hydrides (e.g., H_3S, LaH_{10}) have demonstrated record-high critical temperatures (T_c) under extreme conditions, reigniting interest in optimizing phonon-driven superconductivity through tailored lattice dynamics. Computational tools like DFT and ML are now highly expected to aid in screening and designing novel superconductors with favorable phonon properties,bridging quantum interactions and macroscopic superconducting behavior for next-generation energy technologies.

Phonon-assisted superionic conductors are materials where lattice vibrations (phonons) enable exceptionally fast ion diffusion through a solid (anions) framework, mimicking liquid-like ionic mobility. In these systems, phonons lower energy barriers for ion (cations) hopping by dynamically distorting the lattice, creating transient pathways for ions to migrate. This mechanism is critical in high-performance solid electrolytes (e.g., lithium garnets) used in batteries and fuel cells. Studies from 2016–2025 have leveraged computational tools like AIMD and ML to design materials with optimized phonon spectra and ion-phonon coupling, aiming to enhance ionic conductivity while retaining structural stability. Such advancements promise safer, more efficient energy storage technologies by replacing flammable liquid electrolytes with robust phonon-engineered solids.

Advancements in AI/ML algorithms for phonon research promise to unlock unprecedented insights into thermal transport across the material universe. By integrating generative design, uMLPs, and multi-resolution data strategies, these tools will enable rational

engineering of materials with tailored thermal properties—ushering in a new era of innovation for energy, electronics, and quantum technologies.

5. CONCLUSION

ML is rapidly reshaping the prediction of κ_L in inorganic crystals by offering scalable, data-driven alternatives to conventional first principles and MD approaches. In this review, we first outlined the foundations of prediction using classical MD and the DFT+BTE framework, establishing the baseline for accuracy and computational cost. We then examined progress in AI/ML-based methods since 2016, categorizing them into direct models, which map crystal structure descriptors or learned representations directly to κ_L, and indirect models, which accelerate intermediate steps such as force and IFC calculations through MLPs. We summarized advances in descriptor engineering, GNNs, and ML force fields, and discussed emerging strategies such as multi-fidelity learning, transfer learning, and inverse design. Key challenges remain, including limited high-quality phonon datasets, generalization to unexplored chemistries, and rigorous uncertainty quantification, yet the trajectory of progress is clear. Looking ahead, the integration of ML with physics-based constraints, active learning, and experimental feedback loops is poised to deliver accurate, interpretable, and high-throughput κ_L prediction tools. Such approaches will be instrumental in accelerating the discovery and optimization of materials for thermoelectric, thermal barrier coatings, and solid-state ionic conductors, bridging the gap between theoretical design and real-world applications.

ACKNOWLEDGMENTS

This work was supported in part by the NSF (Award Nos. 2030128, 2110033, 2311202, and 2320292), SC EPSCoR/IDeA Program under NSF OIA-1655740 (23-GC01).

DATA AVAILABILITY

There is no new data produced in this article.

REFERENCES

1. Wei, Y., Liu, Z., and Qin, G., Prediction Methods for Phonon Transport Properties of Inorganic Crystals: From Traditional Approaches to Artificial Intelligence, *Nanoscale Horiz.*, vol. **10**, pp. 230–257, 2025.
2. Zeng, Z., Shen, X., Cheng, R., Perez, O., Ouyang, N., Fan, Z., Lemoine, P., Raveau, B., Guilmeau, E., and Chen, Y., Pushing Thermal Conductivity to Its Lower Limit in Crystals with Simple Structures, *Nat. Commun.*, vol. **15**, no. 1, p. 3007, 2024.
3. McClure, Z.D., Reeve, S.T., and Strachan, A., Role of Electronic Thermal Transport in Amorphous Metal Recrystallization: A Molecular Dynamics Study, *J. Chem. Phys.*, vol. **149**, no. 6, p. 064502, 2018.

4. Smith, M.F. and Whitley, J.B., Properties of Materials, in *Physics of Plasma-Wall Interactions in Controlled Fusion*, Boston: Springer US, pp. 539–605, 1986.
5. Liao, B., Qiu, B., Zhou, J., Huberman, S., Esfarjani, K., and Chen, G., Significant Reduction of Lattice Thermal Conductivity by the Electron-Phonon Interaction in Silicon with High Carrier Concentrations: A First-Principles Study, *Phys. Rev. Lett.*, vol. **114**, no. 11, p. 115901, 2015.
6. Protik, N.H. and Kozinsky, B., Electron-Phonon Drag Enhancement of Transport Properties from a Fully Coupled *Ab Initio* Boltzmann Formalism, *Phys. Rev. B*, vol. **102**, no. 24, p. 245202, 2020.
7. Chen, Z., Zhang, X., and Pei, Y., Manipulation of Phonon Transport in Thermoelectrics, *Adv. Mater.*, vol. **30**, no. 17, p. 1705617, 2018.
8. Kim, W., Strategies for Engineering Phonon Transport in Thermoelectrics, *J. Mater. Chem. C*, vol. **3**, no. 40, pp. 10336–10348, 2015.
9. Yin, Z.P., Kutepov, A., and Kotliar, G., Correlation-Enhanced Electron-Phonon Coupling: Applications of GW and Screened Hybrid Functional to Bismuthates, Chloronitrides, and Other High-T_c Superconductors, *Phys. Rev. X*, vol. **3**, no. 2, p. 021011, 2013.
10. Bardeen, J., Electron-Phonon Interactions and Superconductivity, *Science*, vol. **181**, no. 4106, pp. 1209–1214, 1973.
11. Chen, G., Phonon Transport in Low-Dimensional Structures, *Semicond. Semimetals*, vol. **71**, pp. 203–259, 2001.
12. Khan, S., Shi, X., Feser, J., and Wilson, R., Thermal Conductance of Interfaces between Titanium Nitride and Group IV Semiconductors at High Temperatures, *Appl. Phys. Lett.*, vol. **125**, no. 4, p. 041601, 2024.
13. Varnavides, G., Jermyn, A.S., Anikeeva, P., and Narang, P., Nonequilibrium Phonon Transport across Nanoscale Interfaces, *Phys. Rev. B*, vol. **100**, no. 11, p. 115402, 2019.
14. Kefayati, A., Bird, J.P., and Perebeinos, V., Detection of Single Phonons via Phonon Drag in Two-Dimensional Materials, *Phys. Rev. B*, vol. **106**, no. 15, p. 155415, 2022.
15. Wang, Q., Wang, C., Chi, C., Ouyang, N., Guo, R., Yang, N., and Chen, Y., Phonon Transport in Freestanding $SrTiO_3$ Down to the Monolayer Limit, *Phys. Rev. B*, vol. **108**, no. 11, p. 115435, 2023.
16. Piyathilaka, H.P., Sooriyagoda, R., Whiteside, V.R., Mishima, T.D., Santos, M.B., Sellers, I.R., and Bristow, A.D., Hot-Carrier Dynamics and Transport in III–V Heterostructures for Photovoltaic Applications, *J. Photon. Energy*, vol. **12**, no. 3, p. 032209, 2022.
17. Snyder, G.J. and Snyder, A.H., Figure of Merit ZT of a Thermoelectric Device Defined from Materials Properties, *Energy Environ. Sci.*, vol. **10**, no. 11, pp. 2280–2283, 2017.
18. Ma, Z., Wei, J., Song, P., Zhang, M., Yang, L., Ma, J., Liu, W., Yang, F., and Wang, X., Review of Experimental Approaches for Improving zT of Thermoelectric Materials, *Mater. Sci. Semicond. Process.*, vol. **121**, p. 105303, 2021.
19. Zhang, X. and Zhao, L.D., Thermoelectric Materials: Energy Conversion between Heat and Electricity, *J. Materiomics*, vol. **1**, no. 2, pp. 92–105, 2015.
20. Boeri, L., Hennig, R., Hirschfeld, P., Profeta, G., Sanna, A., Zurek, E., Pickett, W.E., Amsler, M., Dias, R., Eremets, M.I., Heil, C., Hemley, R.J., Liu, H., Ma, Y., Pierleoni, C., Kolmogorov, A.N., Rybin, N., Novoselov, D., Anisimov, V., Oganov, A.R., Pickard, C.J.,

Bi, T., Arita, R., Errea, I., Pellegrini, C., Requist, R., Gross, E.K.U., Margine, E.R., Xie, S.R., Quan, Y., Hire, A., Fanfarillo, L., Stewart, G.R., Hamlin, J.J., Stanev, V., Gonnelli, R.S., Piatti, E., Romanin, D., Daghero, D., and Valenti, R., The 2021 Room-Temperature Superconductivity Roadmap, *J. Phys. Condens. Matter*, vol. **34**, no. 18, p. 183002, 2022.

21. Jurgilaitis, A., Enquist, H., Andreasson, B.P., Persson, A.I.H., Borg, B.M., Caroff, P., Dick, K.A., Harb, M., Linke, H., Nüske, R., Wernersson, L.E., and Larsson, J., Time-Resolved X-Ray Diffraction Investigation of the Modified Phonon Dispersion in InSb Nanowires, *Nano Lett.*, vol. **14**, no. 2, pp. 541–546, 2014.

22. Sheng, T., Liu, X.Z., Qian, L.X., Xu, B., and Zhang, Y.Y., Photoelectric Properties of β-Ga_2O_3 Thin Films Annealed at Different Conditions, *Rare Metals*, vol. **41**, no. 4, pp. 1375–1379, 2022.

23. Novikov, V.V., Mitroshenkov, N.V., Morozov, A.V., Plokhikh, I.V., Pfitzner, A., and Shevelkov, A.V., X-Ray Diffraction Study of Phonon and Magnon Properties of $Eu_2Cu_6P_5$ Ferromagnet, *J. Magn. Magn. Mater.*, vol. **514**, p. 167271, 2020.

24. Chabal, Y.J., Surface Infrared Spectroscopy, *Surf. Sci. Rep.*, vol. **8**, nos. 5-7, pp. 211–357, 1988.

25. Damascelli, A., Schulte, K., van der Marel, D., and Menovsky, A.A., Infrared Spectroscopic Study of Phonons Coupled to Charge Excitations in FeSi, *Phys. Rev. B*, vol. **55**, no. 8, pp. R4863–R4866, 1997.

26. Kuz'menko, A.B., van der Marel, D., van Bentum, P.J.M., Tishchenko, E.A., Presura, C., and Bush, A.A., Infrared Spectroscopic Study of CuO: Signatures of Strong Spin-Phonon Interaction and Structural Distortion, *Phys. Rev. B*, vol. **63**, no. 9, p. 094303, 2001.

27. Hart, T.R., Aggarwal, R.L., and Lax, B., Temperature Dependence of Raman Scattering in Silicon, *Phys. Rev. B*, vol. **1**, no. 2, pp. 638–642, 1970.

28. Rice, C., Young, R.J., Zan, R., Bangert, U., Wolverson, D., Georgiou, T., Jalil, R., and Novoselov, K.S., Raman-Scattering Measurements and First-Principles Calculations of Strain-Induced Phonon Shifts in Monolayer MoS_2, *Phys. Rev. B*, vol. **87**, no. 8, p. 081307, 2013.

29. Mishra, S.K., Gupta, M.K., Mittal, R., Zbiri, M., Rols, S., Schober, H., and Chaplot, S.L., Phonon Dynamics and Inelastic Neutron Scattering of Sodium Niobate, *Phys. Rev. B*, vol. **89**, no. 18, p. 184303, 2014.

30. Pintschovius, L., Electron–Phonon Coupling Effects Explored by Inelastic Neutron Scattering, *Phys. Status Solidi B*, vol. **242**, no. 1, pp. 30–50, 2005.

31. Yuan, C., Hanus, R., and Graham, S., A Review of Thermoreflectance Techniques for Characterizing Wide Bandgap Semiconductors' Thermal Properties and Devices' Temperatures, *J. Appl. Phys.*, vol. **132**, no. 22, p. 220701, 2022.

32. Koh, Y.K., Lyons, A.S., Bae, M.H., Huang, B., Dorgan, V.E., Cahill, D.G., and Pop, E., Role of Remote Interfacial Phonon (RIP) Scattering in Heat Transport Across Graphene/SiO_2 Interfaces, *Nano Lett.*, vol. **16**, no. 10, pp. 6014–6020, 2016.

33. Wei, Y., Liu, Z., and Qin, G., Prediction Methods for Phonon Transport Properties of Inorganic Crystals: From Traditional Approaches to Artificial Intelligence, *Nanoscale Horiz.*, vol. **10**, pp. 230–257, 2025.

34. Zheng, J., Shi, D., Yang, Y., Lin, C., Huang, H., Guo, R., and Huang, B., Anharmonicity-Induced Phonon Hardening and Phonon Transport Enhancement in Crystalline Perovskite $BaZrO_3$, *Phys. Rev. B*, vol. **105**, no. 22, p. 224303, 2022.

35. Parks, H.L., Kim, H.Y., Viswanathan, V., and McGaughey, A.J.H., Uncertainty Quantification in First-Principles Predictions of Phonon Properties and Lattice Thermal Conductivity, *Phys. Rev. Mater.*, vol. **4**, no. 8, p. 083805, 2020.

36. Peng, J., Deskins, W.R., and El-Azab, A., Monte-Carlo Modeling of Phonon Thermal Transport Using DFT-Based Anisotropic Dispersion Relations over the Full Brillouin Zone, *Comput. Mater. Sci.*, vol. **211**, p. 111528, 2022.

37. Hu, M., Unleashing the Power of Artificial Intelligence in Phonon Thermal Transport: Current Challenges and Prospects, *J. Appl. Phys.*, vol. **135**, no. 17, p. 170904, 2024.

38. McGaughey, A. and Kaviany, M., Phonon Transport in Molecular Dynamics Simulations: Formulation and Thermal Conductivity Prediction, *Adv. Heat Transf.*, vol. **39**, pp. 169–255, 2006.

39. McGaughey, A.J.H. and Larkin, J.M., Predicting Phonon Properties from Equilibrium Molecular Dynamics Simulations, *Annu. Rev. Heat Transf.*, vol. **17**, pp. 49–87, 2014.

40. Turney, J.E., Landry, E.S., McGaughey, A.J.H., and Amon, C.H., Predicting Phonon Properties and Thermal Conductivity from Anharmonic Lattice Dynamics Calculations and Molecular Dynamics Simulations, *Phys. Rev. B*, vol. **79**, no. 6, p. 064301, 2009.

41. Carrete, J., Li, W., Mingo, N., Wang, S., and Curtarolo, S., Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling, *Phys. Rev. X*, vol. **4**, no. 1, p. 011019, 2014.

42. Butler, K.T., Davies, D.W., Cartwright, H., Isayev, O., and Walsh, A., Machine Learning for Molecular and Materials Science, *Nature*, vol. **559**, no. 7715, pp. 547–555, 2018.

43. Schmidt, J., Marques, M.R.G., Botti, S., and Marques, M.A.L., Recent Advances and Applications of Machine Learning in Solid-State Materials Science, *NPJ Comput. Mater.*, vol. **5**, no. 1, p. 83, 2019.

44. Wang, A.Y.T., Murdock, R.J., Kauwe, S.K., Oliynyk, A.O., Gurlo, A., Brgoch, J., Persson, K.A., and Sparks, T.D., Machine Learning for Materials Scientists: An Introductory Guide toward Best Practices, *Chem. Mater.*, vol. **32**, no. 12, pp. 4954–4965, 2020.

45. Huo, S., Zhang, S., Wu, Q., and Zhang, X., Feature-Assisted Machine Learning for Predicting Band Gaps of Binary Semiconductors, *Nanomaterials*, vol. **14**, no. 5, p. 445, 2024.

46. Isayev, O., Oses, C., Toher, C., Gossett, E., Curtarolo, S., and Tropsha, A., Universal Fragment Descriptors for Predicting Properties of Inorganic Crystals, *Nat. Commun.*, vol. **8**, no. 1, p. 15679, 2017.

47. Zuo, Y., Chen, C., Li, X., Deng, Z., Chen, Y., Behler, J., Csányi, G., Shapeev, A.V., Thompson, A.P., Wood, M.A., and Ong, S.P., Performance and Cost Assessment of Machine Learning Interatomic Potentials, *J. Phys. Chem. A*, vol. **124**, no. 4, pp. 731–745, 2020.

48. Ju, S. and Shiomi, J., Materials Informatics for Heat Transfer: Recent Progresses and Perspectives, *Nanoscale Microscale Thermophys. Eng.*, vol. **23**, no. 2, pp. 157–172, 2019.

49. Qian, X. and Yang, R., Machine Learning for Predicting Thermal Transport Properties of Solids, *Mater. Sci. Eng. R Rep.*, vol. **146**, p. 100642, 2021.

50. Zou, J., Hirokawa, T., An, J., Huang, L., and Camm, J., Recent Advances in the Applications

of Machine Learning Methods for Heat Exchanger Modeling—A Review, *Front. Energy Res.*, vol. **11**, p. 1294531, 2023.

51. Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G.E., Physics-Informed Neural Networks for Heat Transfer Problems, *J. Heat Transf.*, vol. **143**, no. 6, p. 060801, 2021.
52. Hohenberg, P. and Kohn, W., Inhomogeneous Electron Gas, *Phys. Rev.*, vol. **136**, no. 3B, pp. B864–B871, 1964.
53. Kohn, W. and Sham, L.J., Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.*, vol. **140**, no. 4A, pp. A1133–A1138, 1965.
54. Perdew, J.P., Burke, K., and Ernzerhof, M., Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, vol. **77**, no. 18, pp. 3865–3868, 1996.
55. Adamo, C. and Barone, V., Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model, *J. Chem. Phys.*, vol. **110**, no. 13, pp. 6158–6170, 1999.
56. Baroni, S., de Gironcoli, S., Dal Corso, A., and Giannozzi, P., Phonons and Related Crystal Properties from Density-Functional Perturbation Theory, *Rev. Mod. Phys.*, vol. **73**, no. 2, pp. 515–562, 2001.
57. Hafner, J., *Ab-Initio* Simulations of Materials Using VASP: Density-Functional Theory and Beyond, *J. Comput. Chem.*, vol. **29**, no. 13, pp. 2044–2078, 2008.
58. Togo, A. and Tanaka, I., First Principles Phonon Calculations in Materials Science, *Scr. Mater.*, vol. **108**, pp. 1–5, 2015.
59. Gonze, X. and Lee, C., Dynamical Matrices, Born Effective Charges, Dielectric Permittivity Tensors, and Interatomic Force Constants from Density-Functional Perturbation Theory, *Phys. Rev. B*, vol. **55**, no. 16, pp. 10355–10368, 1997.
60. Esfarjani, K. and Stokes, H.T., Method to Extract Anharmonic Force Constants from First Principles Calculations, *Phys. Rev. B*, vol. **77**, no. 14, p. 144112, 2008.
61. Broido, D.A., Malorny, M., Birner, G., Mingo, N., and Stewart, D.A., Intrinsic Lattice Thermal Conductivity of Semiconductors from First Principles, *Appl. Phys. Lett.*, vol. **91**, no. 23, p. 231922, 2007.
62. Li, W., Carrete, J., Katcho, N.A., and Mingo, N., ShengBTE: A Solver of the Boltzmann Transport Equation for Phonons, *Comput. Phys. Commun.*, vol. **185**, no. 6, pp. 1747–1758, 2014.
63. Kresse, G. and Furthmüller, J., Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set, *Phys. Rev. B*, vol. **54**, no. 16, pp. 11169–11186, 1996.
64. Kresse, G. and Hafner, J., *Ab Initio* Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium, *Phys. Rev. B*, vol. **49**, no. 20, pp. 14251–14269, 1994.
65. Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G.L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A.P., Smogunov, A., Umari, P., and Wentzcovitch, R.M., QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials, *J. Phys. Condens. Matter*, vol. **21**, no. 39, p. 395502, 2009.

66. Gonze, X., Amadon, B., Anglade, P.M., Beuken, J.M., Bottin, F., Boulanger, P., Bruneval, F., Caliste, D., Caracas, R., Côté, M., Deutsch, T., Genovese, L., Ghosez, Ph. Giantomassi, M., Goedecker, S., Hamann, D.R., Hermet, P., Jollet, F., Jomard, G., Leroux, S., Mancini, M., Mazevet, S., Oliveira, M.J.T., Onida, G., Pouillon, Y., Rangel, T., Rignanese, G.-M., Sangalli, D., Shaltaf, R., Torrent, M., Verstraete, M.J., Zerah, G., and Zwanziger, J.W., ABINIT: First-Principles Approach to Material and Nanosystem Properties, *Comput. Phys. Commun.*, vol. **180**, no. 12, pp. 2582–2615, 2009.

67. Gonze, X., Jollet, F., Abreu Araujo, F., Adams, D., Amadon, B., Applencourt, T., Audouze, C., Beuken, J.-M., Bieder, J., Bokhanchuk, A., Bousquet, E., Bruneval, F., Caliste, D., Côté, M., Dahm, F., Da Pieve, F., Delaveau, M., Di Gennaro, M., Dorado, B., Espejo, C., Geneste, G., Genovese, L., Gerossier, A., Giantomassi, M., Gillet, Y., Hamann, D.R., He, L., Jomard, G., Laflamme Janssen, J., Le Roux, S., Levitt, A., Lherbier, A., Liu, F., Lukaèeviæ, I., Martin, A., Martins, C., Oliveira, M.J.T., Poncé, S., Pouillon, Y., Rangel, T., Rignanese, G.-M., Romero, A.H., Rousseau, B., Rubel, O., Shukri, A.A., Stankovski, M., Torrent, M., Van Setten, M.J., Van Troeye, B., Verstraete, M.J., Waroquiers, D., Wiktor, J., Xu, B., Zhou, A., and Zwanziger, J.W., Recent Developments in the ABINIT Software Package, *Comput. Phys. Commun.*, vol. **205**, pp. 106–131, 2016.

68. Togo, A., Chaput, L., and Tanaka, I., Distributions of Phonon Lifetimes in Brillouin Zones, *Phys. Rev. B*, vol. **91**, no. 9, p. 094306, 2015.

69. Zhou, F., Sadigh, B., Åberg, D., Xia, Y., and Ozoliņš, V., Compressive Sensing Lattice Dynamics. II. Efficient Phonon Calculations and Long-Range Interactions, *Phys. Rev. B*, vol. **100**, no. 18, p. 184309, 2019.

70. Zhou, F., Nielson, W., Xia, Y., and Ozoliņš, V., Compressive Sensing Lattice Dynamics. I. General Formalism, *Phys. Rev. B*, vol. **100**, no. 18, p. 184308, 2019.

71. Srivastava, G.P., *The Physics of Phonons*, Boca Raton, FL: CRC Press, 2022.

72. Ziman, J.M., *Electrons and Phonons*, Oxford, UK: Oxford University Press, 2001.

73. Wang, H., Qin, G., Qin, Z., Li, G., Wang, Q., and Hu, M., Lone-Pair Electrons Do Not Necessarily Lead to Low Lattice Thermal Conductivity: An Exception of Two-Dimensional Penta-CN_2, *J. Phys. Chem. Lett.*, vol. **9**, no. 10, pp. 2474–2483, 2018.

74. Ouyang, T. and Hu, M., Thermal Transport and Thermoelectric Properties of Beta-Graphyne Nanostructures, *Nanotechnology*, vol. **25**, no. 24, p. 245401, 2014.

75. Yue, S.Y., Ouyang, T., and Hu, M., Diameter Dependence of Lattice Thermal Conductivity of Single-Walled Carbon Nanotubes: Study from *Ab Initio*, *Sci. Rep.*, vol. **5**, no. 1, p. 15440, 2015.

76. Qin, G., Qin, Z., Wang, H., and Hu, M., On the Diversity in the Thermal Transport Properties of Graphene: A First-Principles-Benchmark Study Testing Different Exchange-Correlation Functionals, *Comput. Mater. Sci.*, vol. **151**, pp. 153–159, 2018.

77. Qin, G., Wang, H., Zhang, L., Qin, Z., and Hu, M., Giant Effect of Spin–Lattice Coupling on the Thermal Transport in Two-Dimensional Ferromagnetic CrI_3, *J. Mater. Chem. C*, vol. **8**, no. 10, pp. 3520–3526, 2020.

78. Lindsay, L., Broido, D.A., and Reinecke, T.L., First-Principles Determination of Ultrahigh Thermal Conductivity of Boron Arsenide: A Competitor for Diamond?, *Phys. Rev. Lett.*, vol. **111**, no. 2, p. 025901, 2013.

79. Ravichandran, N.K. and Broido, D., Phonon-Phonon Interactions in Strongly Bonded Solids: Selection Rules and Higher-Order Processes, *Phys. Rev. X*, vol. **10**, no. 2, p. 021063, 2020.

80. Cui, Y., Qin, Z., Wu, H., Li, M., and Hu, Y., Flexible Thermal Interface Based on Self-Assembled Boron Arsenide for High-Performance Thermal Management, *Nat. Commun.*, vol. **12**, no. 1, p. 1284, 2021.

81. Kang, J.S., Li, M., Wu, H., Nguyen, H., and Hu, Y., Experimental Observation of High Thermal Conductivity in Boron Arsenide, *Science*, vol. **361**, no. 6402, pp. 575–578, 2018.

82. Zhao, L.D., Lo, S.H., Zhang, Y., Sun, H., Tan, G., Uher, C., Wolverton, C., Dravid, V.P., and Kanatzidis, M.G., Ultralow Thermal Conductivity and High Thermoelectric Figure of Merit in SnSe Crystals, *Nature*, vol. **508**, no. 7496, pp. 373–377, 2014.

83. Carrete, J., Mingo, N., and Curtarolo, S., Low Thermal Conductivity and Triaxial Phononic Anisotropy of SnSe, *Appl. Phys. Lett.*, vol. **105**, no. 10, p. 101907, 2014.

84. Ryu, B., Chung, J., Choi, E.-A., Ziolkowski, P., Muller, E., and Park, S., Counterintuitive Example on Relation between ZT and Thermoelectric Efficiency, *Appl. Phys. Lett.*, vol. **116,** no. 19, p. 193903, 2020.

85. Wolf, M., Hinterding, R., and Feldhoff, A., High Power Factor vs. High zT—A Review of Thermoelectric Materials for High-Temperature Application, *Entropy*, vol. **21**, no. 11, p. 1058, 2019.

86. Wang, X., Gao, Z., Zhu, G., Ren, J., Hu, L., Sun, J., Ding, X., Xia, Y., and Li, B., Role of High-Order Anharmonicity and Off-Diagonal Terms in Thermal Conductivity: A Case Study of Multiphase $CsPbBr_3$, *Phys. Rev. B*, vol. **107**, no. 21, p. 214308, 2023.

87. Zhang, C., Sun, J., Shen, Y., Kang, W., and Wang, Q., Effect of High Order Phonon Scattering on the Thermal Conductivity and Its Response to Strain of a Penta-NiN_2 Sheet, *J. Phys. Chem. Lett.*, vol. **13**, no. 25, pp. 5734–5741, 2022.

88. Chaudhuri, S., Bhattacharya, A., Das, A.K., Das, G.P., and Dev, B.N., Understanding the Role of Four-Phonon Scattering in the Lattice Thermal Transport of Monolayer MoS_2, *Phys. Rev. B*, vol. **109**, no. 23, p. 235424, 2024.

89. Wang, X., Feng, M., Xia, Y., Sun, J., Ding, X., Li, B., and Gao, Z., Revisiting Lattice Thermal Conductivity of CsCl: The Crucial Role of Quartic Anharmonicity, *Appl. Phys. Lett.*, vol. **124**, no. 17, p. 170501, 2024.

90. Feng, T. and Ruan, X., Quantum Mechanical Prediction of Four-Phonon Scattering Rates and Reduced Thermal Conductivity of Solids, *Phys. Rev. B*, vol. **93**, no. 4, p. 045202, 2016.

91. Castellano, A., Alvarinhaus Batista, J.P., Hellman, O., and Verstraete, M.J., Mode-Coupling Formulation of Heat Transport in Anharmonic Materials, *Phys. Rev. B*, vol. **111**, no. 9, p. 094306, 2025.

92. Liang, Z. and Hu, M., Tutorial: Determination of Thermal Boundary Resistance by Molecular Dynamics Simulations, *J. Appl. Phys.*, vol. **123**, no. 19, p. 191101, 2018.

93. Zhang, X., Hu, M., and Poulikakos, D., A Low-Frequency Wave Motion Mechanism Enables Efficient Energy Transport in Carbon Nanotubes at High Heat Fluxes, *Nano Lett.*, vol. **12**, no. 7, pp. 3410–3416, 2012.

94. Bao, H., Shao, C., Luo, S., and Hu, M., Enhancement of Interfacial Thermal Transport by Carbon Nanotube-Graphene Junction, *J. Appl. Phys.*, vol. **115**, no. 5, p. 053524, 2014.

95. Zhou, Y., Xiong, S., Zhang, X., Volz, S., and Hu, M., Thermal Transport Crossover from

Crystalline to Partial-Crystalline Partial-Liquid State, *Nat. Commun.*, vol. **9**, no. 1, p. 4712, 2018.

96. Fang, F. and Zhu, P., Molecular Dynamics, in *CIRP Encyclopedia of Production Engineering*, Berlin: Springer, pp. 1236–1239, 2019.
97. Kedharnath, A., Kapoor, R., and Sarkar, A., Classical Molecular Dynamics Simulations of the Deformation of Metals under Uniaxial Monotonic Loading: A Review, *Comput. Struct.*, vol. **254**, p. 106614, 2021.
98. Stillinger, F.H. and Weber, T.A., Computer Simulation of Local Order in Condensed Phases of Silicon, *Phys. Rev. B*, vol. **31**, no. 8, pp. 5262–5271, 1985.
99. Likos, C.N., Effective Interactions in Soft Condensed Matter Physics, *Phys. Rep.*, vol. **348**, nos. 4-5, pp. 267–439, 2001.
100. Weeks, J.D., Chandler, D., and Andersen, H.C., Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids, *J. Chem. Phys.*, vol. **54**, no. 12, pp. 5237–5247, 1971.
101. Rao, P.S., Anandatheertha, S., Naik, G.N., and Gopalakrishnan, S., Estimation of Mechanical Properties of Single Wall Carbon Nanotubes Using Molecular Mechanics Approach, *Sadhana*, vol. **40**, no. 4, pp. 1301–1311, 2015.
102. Belytschko, T., Xiao, S.P., Schatz, G.C., and Ruoff, R.S., Atomistic Simulations of Nanotube Fracture, *Phys. Rev. B*, vol. **65**, no. 23, p. 235430, 2002.
103. Tersoff, J., Empirical Interatomic Potential for Silicon with Improved Elastic Properties, *Phys. Rev. B*, vol. **38**, no. 14, pp. 9902–9905, 1988.
104. Batatia, I., Kovacs, D.P., Simm, G.N.C., Ortner, C., and Csanyi, G., MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, *Adv. Neural Inf. Process. Syst.*, vol. **35**, pp. 11423–11436, 2022.
105. Ozsipahi, M., Jean, S., Beskok, A., and Wilson, A.A., Molecular Dynamics Simulation of Thermal Conductivity of GaN, *Int. Commun. Heat Mass Transf.*, vol. **163**, p. 108658, 2025.
106. Müller-Plathe, F., A Simple Nonequilibrium Molecular Dynamics Method for Calculating the Thermal Conductivity, *J. Chem. Phys.*, vol. **106**, no. 14, pp. 6082–6085, 1997.
107. Plimpton, S., Fast Parallel Algorithms for Short-Range Molecular Dynamics, *J. Comput. Phys.*, vol. **117**, no. 1, pp. 1–19, 1995.
108. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindahl, E., GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers, *SoftwareX*, vols. **1–2**, pp. 19–25, 2015.
109. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K., Scalable Molecular Dynamics with NAMD, *J. Comput. Chem.*, vol. **26**, no. 16, pp. 1781–1802, 2005.
110. Smith, W. and Forester, T.R., DL_POLY_2.0: A General-Purpose Parallel Molecular Dynamics Simulation Package, *J. Mol. Graph.*, vol. **14**, no. 3, pp. 136–141, 1996.
111. Anderson, J.A., Glaser, J., and Glotzer, S.C., HOOMD-Blue: A Python Package for High-Performance Molecular Dynamics and Hard Particle Monte Carlo Simulations, *Comput. Mater. Sci.*, vol. **173**, p. 109363, 2020.
112. Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A.,

Simmerling, C., Wang, B., and Woods, R.J., The Amber Biomolecular Simulation Programs, *J. Comput. Chem.*, vol. **26**, no. 16, pp. 1668–1688, 2005.

113. Daw, M.S. and Baskes, M.I., Embedded-Atom Method: Derivation and Application to Impurities, Surfaces, and Other Defects in Metals, *Phys. Rev. B*, vol. **29**, no. 12, pp. 6443–6453, 1984.
114. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M., CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comput. Chem.*, vol. **4**, no. 2, pp. 187–217, 1983.
115. Salaway, R.N. and Zhigilei, L.V., Molecular Dynamics Simulations of Thermal Conductivity of Carbon Nanotubes: Resolving the Effects of Computational Parameters, *Int. J. Heat Mass Transf.*, vol. **70**, pp. 954–964, 2014.
116. Arya, G., Chang, H.C., and Maginn, E.J., A Critical Comparison of Equilibrium, Non-Equilibrium and Boundary-Driven Molecular Dynamics Techniques for Studying Transport in Microporous Materials, *J. Chem. Phys.*, vol. **115**, no. 17, pp. 8112–8124, 2001.
117. Chen, J., Zhang, G., and Li, B., A Universal Gauge for Thermal Conductivity of Silicon Nanowires with Different Cross Sectional Geometries, *J. Chem. Phys.*, vol. **135**, no. 20, p. 204705, 2011.
118. Liang, Z., Jain, A., McGaughey, A.J.H., and Keblinski, P., Molecular Simulations and Lattice Dynamics Determination of Stillinger-Weber GaN Thermal Conductivity, *J. Appl. Phys.*, vol. **118**, no. 12, p. 125104, 2015.
119. Zhou, X.W., Aubry, S., Jones, R.E., Greenstein, A., and Schelling, P.K., Towards more Accurate Molecular Dynamics Calculation of Thermal Conductivity: Case Study of GaN Bulk Crystals, *Phys. Rev. B*, vol. **79**, no. 11, p. 115201, 2009.
120. Matsubara, H., Kikugawa, G., Ishikiriyama, M., Yamashita, S., and Ohara, T., Equivalence of the EMD- and NEMD-Based Decomposition of Thermal Conductivity into Microscopic Building Blocks, *J. Chem. Phys.*, vol. **147**, no. 11, p. 114104, 2017.
121. Dong, H., Fan, Z., Shi, L., Harju, A., and Ala-Nissila, T., Equivalence of the Equilibrium and the Nonequilibrium Molecular Dynamics Methods for Thermal Conductivity Calculations: From Bulk to Nanowire Silicon, *Phys. Rev. B*, vol. **97**, no. 9, p. 094305, 2018.
122. Kühne, T.D., Second Generation Car–Parrinello Molecular Dynamics, *WIREs Comput. Mol. Sci.*, vol. **4**, no. 4, pp. 391–406, 2014.
123. Ratcliff, L.E., Dawson, W., Fisicaro, G., Caliste, D., Mohr, S., Degomme, A., Videau, B., Cristiglio, V., Stella, M., D'Alessandro, M., Goedecker, S., Nakajima, T., Deutsch, T., and Genovese, L., Flexibilities of Wavelets as a Computational Basis Set for Large-Scale Electronic Structure Calculations, *J. Chem. Phys.*, vol. **152**, no. 19, p. 194110, 2020.
124. Motamarri, P., Das, S., Rudraraju, S., Ghosh, K., Davydov, D., and Gavini, V., DFT-FE—A Massively Parallel Adaptive Finite-Element Code for Large-Scale Density Functional Theory Calculations, *Comput. Phys. Commun.*, vol. **246**, p. 106853, 2020.
125. Behler, J., Perspective: Machine Learning Potentials for Atomistic Simulations, *J. Chem. Phys.*, vol. **145**, no. 17, p. 170901, 2016.
126. Phonondb@Kyoto University, accessed from https://github.com/atztogo/phonondb, 2024.
127. Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., and Persson, K.A., Commentary: The Materials Project: A Materials

Genome Approach to Accelerating Materials Innovation, *APL Mater.*, vol. **1**, no. 1, p. 011002, 2013.

128. Jain, A., Machine Learning in Materials Research: Developments over the Last Decade and Challenges for the Future, *Curr. Opin. Solid State Mater. Sci.*, vol. **33**, p. 101189, 2024.
129. Curtarolo, S., Setyawan, W., Wang, S., Xue, J., Yang, K., Taylor, R.H., Nelson, L.J., Hart, G.L.W., Sanvito, S., Buongiorno-Nardelli, M., Mingo, N., and Levy, O., AFLOWLIB.ORG: A Distributed Materials Properties Repository from High-Throughput *Ab Initio* Calculations, *Comput. Mater. Sci.*, vol. **58**, pp. 227–235, 2012.
130. Saal, J.E., Kirklin, S., Aykol, M., Meredig, B., and Wolverton, C., Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD), *JOM*, vol. **65**, no. 11, pp. 1501–1509, 2013.
131. Draxl, C. and Scheffler, M., Big Data-Driven Materials Science and Its FAIR Data Infrastructure, in *Handbook of Materials Modeling*, Cham: Springer, pp. 1–25, 2019.
132. Smith, J.S., Nebgen, B., Lubbers, N., Isayev, O., and Roitberg, A.E., Less Is More: Sampling Chemical Space with Active Learning, *J. Chem. Phys.*, vol. **148**, no. 24, p. 241733, 2018.
133. Kim, J., Kim, J., Kim, J., Lee, J., Park, Y., Kang, Y., and Han, S., Data-Efficient Multifidelity Training for High-Fidelity Machine Learning Interatomic Potentials, *J. Am. Chem. Soc.*, vol. **147**, no. 1, pp. 1042–1054, 2025.
134. Ghiringhelli, L.M., Vybiral, J., Levchenko, S.V., Draxl, C., and Scheffler, M., Big Data of Materials Science: Critical Role of the Descriptor, *Phys. Rev. Lett.*, vol. **114**, no. 10, p. 105503, 2015.
135. Klochko, L., Togo, A. and Chaput, L., Transfer Learning for Deep Learning-Based Prediction of Lattice Thermal Conductivity, arXiv Preprint arXiv:2411.18259, 2024.
136. Liu, Z., Jiang, M., and Luo, T., Leveraging Low-Fidelity Data to Improve Machine Learning of Sparse High-Fidelity Thermal Conductivity Data via Transfer Learning, *Mater. Today Phys.*, vol. **28**, p. 100868, 2022.
137. Ojih, J., Al-Fahdi, M., Rodriguez, A.D., Choudhary, K., and Hu, M., Efficiently Searching Extreme Mechanical Properties via Boundless Objective-Free Exploration and Minimal First-Principles Calculations, *NPJ Comput. Mater.*, vol. **8**, no. 1, p. 143, 2022.
138. Ojih, J., Al-Fahdi, M., Yao, Y., Hu, J., and Hu, M., Graph Theory and Graph Neural Network Assisted High-Throughput Crystal Structure Prediction and Screening for Energy Conversion and Storage, *J. Mater. Chem. A*, vol. **12**, no. 14, pp. 8502–8515, 2024.
139. Liu, J., Han, S., Cao, G., Zhou, Z., Sheng, C., and Liu, H., A High-Throughput Descriptor for Prediction of Lattice Thermal Conductivity of Half-Heusler Compounds, *J. Phys. D Appl. Phys.*, vol. **53**, no. 31, p. 315301, 2020.
140. Qin, G., Wei, Y., Yu, L., Xu, J., Ojih, J., Rodriguez, A.D., Wang, H., Qin, Z., and Hu, M., Predicting Lattice Thermal Conductivity from Fundamental Material Properties Using Machine Learning Techniques, *J. Mater. Chem. A*, vol. **11**, no. 11, pp. 5801–5810, 2023.
141. Chen, L., Tran, H., Batra, R., Kim, C., and Ramprasad, R., Machine Learning Models for the Lattice Thermal Conductivity Prediction of Inorganic Materials, *Comput. Mater. Sci.*, vol. **170**, p. 109155, 2019.
142. Trans, R., Løvvik, O.M., Tomic, O., and Berland, K., Lattice Thermal Conductivity of Half-Heuslers with Density Functional Theory and Machine Learning: Enhancing Predictivity

by Active Sampling with Principal Component Analysis, *Comput. Mater. Sci.*, vol. **202**, p. 110938, 2022.

143. Xie, T. and Grossman, J.C., Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, *Phys. Rev. Lett.*, vol. **120**, no. 14, p. 145301, 2018.
144. Choudhary, K. and DeCost, B., Atomistic Line Graph Neural Network for Improved Materials Property Predictions, *NPJ Comput. Mater.*, vol. **7**, no. 1, p. 185, 2021.
145. Karamad, M., Magar, R., Shi, Y., Siahrostami, S., Gates, I.D., and Barati Farimani, A., Orbital Graph Convolutional Neural Network for Material Property Prediction, *Phys. Rev. Mater.*, vol. **4**, no. 9, p. 093801, 2020.
146. Omee, S.S., Louis, S.Y., Fu, N., Wei, L., Dey, S., Dong, R., Li, Q., and Hu., J., Scalable Deeper Graph Neural Networks for High-Performance Materials Property Prediction, *Patterns*, vol. **3**, no. 5, p. 100491, 2022.
147. Park, C.W. and Wolverton, C., Developing an Improved Crystal Graph Convolutional Neural Network Framework for Accelerated Materials Discovery, *Phys. Rev. Mater.*, vol. **4**, no. 6, p. 063801, 2020.
148. Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S.P., Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals, *Chem. Mater.*, vol. **31**, no. 9, pp. 3564–3572, 2019.
149. Ojih, J., Rodriguez, A., and Hu, M., Screening Outstanding Mechanical Properties and Low Lattice Thermal Conductivity Using Global Attention Graph Neural Network, *Energy AI*, vol. **14**, p. 100286, 2023.
150. Ojih, J., Shen, C., Rodriguez, A., Zhang, H., Choudhary, K., and Hu, M., High-Throughput Computational Discovery of 3218 Ultralow Thermal Conductivity and Dynamically Stable Materials by Dual Machine Learning Models, *J. Mater. Chem. A*, vol. **11**, no. 44, pp. 24169–24183, 2023.
151. Lundberg, S.M. and Lee, S., A Unified Approach to Interpreting Model Predictions, *Adv. Neural Inf. Process. Syst.*, vol. **30**, pp. 4766–4777, 2017.
152. De Jong, M., Chen, W., Notestine, R., Persson, K., Ceder, G., Jain, A., Asta, M., and Gamst, A., A Statistical Learning Framework for Materials Science: Application to Elastic Moduli of k-Nary Inorganic Polycrystalline Compounds, *Sci. Rep.*, vol. **6**, p. 34256, 2016.
153. Ying, Z., Bourgeois, D., You, J., Zitnik, M., and Leskovec, J., GNNExplainer: Generating Explanations for Graph Neural Networks, *Adv. Neural Inf. Process. Syst.*, vol. **32**, pp. 9240–9251, 2019.
154. Schütt, K.T., Sauceda, H.E., Kindermans, P.J., Tkatchenko, A., and Müller, K.R., SchNet—A Deep Learning Architecture for Molecules and Materials, *J. Chem. Phys.*, vol. **148**, no. 24, p. 241722, 2018.
155. Behler, J. and Parrinello, M., Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.*, vol. **98**, no. 14, p. 146401, 2007.
156. Bartk, A.P., Payne, M.C., Kondor, R., and Csányi, G., Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, *Phys. Rev. Lett.*, vol. **104**, no. 13, p. 136403, 2010.
157. Bartók, A.P., Kondor, R., and Csányi, G., On Representing Chemical Environments, *Phys. Rev. B*, vol. **87**, no. 18, p. 184115, 2013.

158. Shapeev, A.V., Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials, *Multiscale Model. Simul.*, vol. **14**, no. 3, pp. 1153–1173, 2016.

159. Podryabinkin, E.V. and Shapeev, A.V., Active Learning of Linearly Parametrized Interatomic Potentials, *Comput. Mater. Sci.*, vol. **140**, pp. 171–180, 2017.

160. Thompson, A.P., Swiler, L.P., Trott, C.R., Foiles, S.M., and Tucker, G.J., Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials, *J. Comput. Phys.*, vol. **285**, pp. 316–330, 2015.

161. Zhang, L., Han, J., Wang, H., Car, R., and E, W., Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, *Phys. Rev. Lett.*, vol. **120**, no. 14, p. 143001, 2018.

162. Wang, H., Zhang, L., Han, J., and E, W., DeePMD-kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics, *Comput. Phys. Commun.*, vol. **228**, pp. 178–184, 2018.

163. Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J.P., Kornbluth, M., Molinari, N., Smidt, T.E., and Kozinsky, B., E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials, *Nat. Commun.*, vol. **13**, p. 2453, 2022.

164. Drautz, R., Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials, *Phys. Rev. B*, vol. **99**, no. 1, p. 014104, 2019.

165. Unke, O.T., Chmiela, S., Sauceda, H.E., Gastegger, M., Poltavsky, I., Schütt, K.T., Tkatchenko, A., and Muller, K.-R., Machine Learning Force Fields, *Chem. Rev.*, vol. **121**, no. 16, pp. 10142–10186, 2021.

166. Morrow, J.D., Gardner, J.L.A., and Deringer, V.L., How to Validate Machine-Learned Interatomic Potentials, *J. Chem. Phys.*, vol. **158**, no. 12, p. 121501, 2023.

167. Wang, G., Wang, C., Zhang, X., Li, Z., Zhou, J., and Sun, Z., Machine Learning Interatomic Potential: Bridge the Gap between Small-Scale Models and Realistic Device-Scale Simulations, *iScience*, vol. **27**, no. 5, p. 109673, 2024.

168. Mortazavi, B., Podryabinkin, E.V., Novikov, I.S., Rabczuk, T., Zhuang, X., and Shapeev, A.V., Accelerating First-Principles Estimation of Thermal Conductivity by Machine-Learning Interatomic Potentials: A MTP/ShengBTE Solution, *Comput. Phys. Commun.*, vol. **258**, p. 107583, 2021.

169. Verdi, C., Karsai, F., Liu, P., Jinnouchi, R., and Kresse, G., Thermal Transport and Phase Transitions of Zirconia by On-the-Fly Machine-Learned Interatomic Potentials, *NPJ Comput. Mater.*, vol. **7**, p. 156, 2021.

170. Korotaev, P., Novoselov, I., Yanilkin, A., and Shapeev, A., Accessing Thermal Conductivity of Complex Compounds by Machine Learning Interatomic Potentials, *Phys. Rev. B*, vol. **100**, no. 14, p. 144308, 2019.

171. Srivastava, Y. and Jain, A., Accelerating Prediction of Phonon Thermal Conductivity by an Order of Magnitude through Machine Learning Assisted Extraction of Anharmonic Force Constants, *Phys. Rev. B*, vol. **110**, no. 16, p. 165202, 2024.

172. Lu, Y. and Zheng, F., Lattice Thermal Conductivity of $ZrSe_2$ Based on the Anharmonic Phonon Approach and On-the-Fly Machine Learning Force Fields, *Phys. Rev. B*, vol. **109**, no. 1, p. 014305, 2024.

173. Ouyang, Y., Yu, C., He, J., Jiang, P., Ren, W., and Chen, J., Accurate Description of High-Order Phonon Anharmonicity and Lattice Thermal Conductivity from Molecular Dynamics Simulations with Machine Learning Potential, *Phys. Rev. B*, vol. **105**, no. 11, p. 115202, 2022.

174. Guo, Z., Roy Chowdhury, P., Han, Z., Sun, Y., Feng, D., Lin, G., and Ruan, X., Fast and Accurate Machine Learning Prediction of Phonon Scattering Rates and Lattice Thermal Conductivity, *NPJ Comput. Mater.*, vol. **9**, p. 95, 2023.

175. Guo, Z., Han, Z., Feng, D., Lin, G., and Ruan, X., Sampling-Accelerated Prediction of Phonon Scattering Rates for Converged Thermal Conductivity and Radiative Properties, *NPJ Comput. Mater.*, vol. **10**, p. 31, 2024.

176. Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C.J., and Ceder, G., CHGNet as a Pretrained Universal Neural Network Potential for Charge-Informed Atomistic Modelling, *Nat. Mach. Intell.*, vol. **5**, no. 9, pp. 1031–1041, 2023.

177. Rodriguez, A., Liu, Y., and Hu, M., Spatial Density Neural Network Force Fields with First-Principles Level Accuracy and Application to Thermal Transport, *Phys. Rev. B*, vol. **102**, no. 3, p. 035203, 2020.

178. Rodriguez, A., Lin, C., Yang, H., Al-Fahdi, M., Shen, C., Choudhary, K., Zhao, Y., Hu, J., Cao, B., Zhang, H., and Hu, M., Million-Scale Data Integrated Deep Neural Network for Phonon Properties of Heuslers Spanning the Periodic Table, *NPJ Comput. Mater.*, vol. **9**, p. 20, 2023.

179. Rodriguez, A., Lin, C., Shen, C., Yuan, K., Al-Fahdi, M., Zhang, X., Zhang, H., and Hu, M., Unlocking Phonon Properties of a Large and Diverse Set of Cubic Crystals by Indirect Bottom-Up Machine Learning Approach, *Commun. Mater.*, vol. **4**, p. 61, 2023.

180. Chen, C. and Ong, S.P., A Universal Graph Deep Learning Interatomic Potential for the Periodic Table, *Nat. Comput. Sci.*, vol. **2**, no. 11, pp. 718–728, 2022.

181. Liao, Y., Wood, B., Das, A., and Smidt, T., EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations, arXiv Preprint arXiv:2306.12059, 2023.

182. Anam, M.Z., Aghoghovbia, O., Al-Fahdi, M., Kong, L., and Hu, M., A Comprehensive Assessment and Benchmark Study of Large Atomistic Foundation Models for Phonons, arXiv Preprint arXiv:2509.03401, 2025.

183. Loftis, C., Yuan, K., Zhao, Y., Hu, M., and Hu, J., Lattice Thermal Conductivity Prediction Using Symbolic Regression and Machine Learning, *J. Phys. Chem. A*, vol. **125**, no. 1, pp. 435–450, 2021.

184. Al-Fahdi, M., Lin, C., Shen, C., Zhang, H., and Hu, M., Rapid Prediction of Phonon Density of States by Crystal Attention Graph Neural Network and High-Throughput Screening of Candidate Substrates for Wide Bandgap Electronic Cooling, *Mater. Today Phys.*, vol. **50**, p. 101632, 2025.

185. Al-Fahdi, M. and Hu, M., High Throughput Substrate Screening for Interfacial Thermal Management of β-Ga_2O_3 by Deep Convolutional Neural Network, *J. Appl. Phys.*, vol. **135**, p. 20, 2024.