Summer 2023

Development of Atomistic Machine Learning Approaches for Thermal Properties of Multi-Component Solids and Liquids

Alejandro David Rodriguez

DEVELOPMENT OF ATOMISTIC MACHINE LEARNING APPROACHES FOR THERMAL PROPERTIES OF MULTI-COMPONENT SOLIDS AND LIQUIDS

by

Alejandro David Rodriguez

Bachelor of Science
University of South Carolina, 2018

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Mechanical Engineering

College of Engineering and Computing

University of South Carolina

2023

Accepted by:

Ming Hu, Major Professor

Theodore M. Besmann,, Committee Member

Jamil A. Khan, Committee Member

Jianjun Hu, Committee Member

Ann Vail, Dean of the Graduate School

Dedication

To my beloved Kerrigan, for her open heart and unyielding kindness. To my parents, brother, and grandmother, for pushing me beyond what I thought was possible. To my closest friends, for their support along the way. Thank you.

Abstract

Currently, heat transfer in many industries is the limiting factor for innovation, especially in the energy sector. For example, maximizing thermal conductivity of ceramic coatings in power plant devices improves the overall electrical to thermal energy ratio, whereas minimizing thermal conductivity is required for desirable heat-to-electricity conversion in thermoelectric devices. As such, rapid discovery of new materials with extreme thermal conductivity values is quintessential for the near-future deployment of current and developing energy applications.

The vibrational properties of crystalline materials are essential for their ability to conduct heat. Fundamentally, the restorative atomic forces of displaced atoms are sufficient to represent phonons, or quasi-particles of vibrations. As such, accurate *ab initio* calculators such as density functional theory (DFT) are standard for computing atomic forces and subsequent interatomic force constants (IFCs) for phonon properties, such as thermodynamic stability and lattice thermal conductivity (LTC). However, depending on the crystal symmetry and the order of the interatomic force constants, several tens to hundreds of simulations containing displaced atoms are required for a single crystalline material. In the context of high-throughput prediction, the costly and time-consuming nature of DFT renders this infeasible when facing tens of thousands of new materials for prediction.

This dissertation focuses on the development, application, and demonstration of machine learning-based models for rapid virtual evaluation and screening of materials for thermal applications. As an atomistic or “bottoms-up” approach, the Spatial Density Neural Network Force Field (SDNNFF) model proposed here predicts atomic forces which may be used in standard phonon calculators and Boltzmann Transport Equation (BTE) codes at significantly faster speeds than DFT. Specifically, atomic forces from reliable quantum mechanical simulations are trained in neural networks, demonstrating a $< 10^{-2}$ eV/Å force accuracy which is on the similar level as pseudopotentials. Additionally, the speed of the SDNNFF relies on the atomic or input descriptors, providing at minimum a 10^3 speed-up for systems containing > 100 atoms in comparison to first-principles DFT. Correspondingly, high-throughput thermal conductivity calculations are made possible for databases containing $> 10^5$ theoretical materials.

To demonstrate database evaluation of LTC, the SDNNFF inputs are modified to accommodate an arbitrary number of elements, which is trained with a dataset spanning up to 63 elements and 16 structural prototypes in a single model. The abundant dataset size of 3×10^7 atomic environments for training is made possible by the training on atomic forces, yielding more information per costly DFT run. Additionally, active learning techniques provide iterative improvement of the model with little to no human intervention. In total, 88,597 materials are evaluated for their thermodynamic stability, whereby the 34,432 remaining materials are queried for lattice thermal conductivity values. Other

exotic properties, such as Weyl points, p-d orbital hybridization, and phonon hydrodynamics, are made possible from the predicted atomic forces. The properties of materials found here are also verified with existing experimental and computational works with agreeable accuracy.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Historically, breakthroughs in materials discovery dictate the very limits of engineering technology. This is especially pronounced in the energy industry, where material properties are major contributors to the performance of electrical generation processes. One application of recent interest is thermoelectric (TE) generation, whereby electrical current is generated by a temperature gradient [1]. Excess heat is generated in many industrial and commercial devices, such as automobiles. On average, more than 60% of energy production is lost as waste heat across various everyday technological applications [2]. In an effort to recover waste heat as productive electrical energy, TE devices are proposed as a partial solution to the increasingly concerning energy crisis. The heat-to-electrical conversion performance is primarily dictated by the intrinsic properties of the material, including the Seebeck coefficient S, electrical conductivity σ, and thermal conductivity κ. Specifically, the ZT score or dimensionless figure of merit is defined as

$$ZT = \frac{S^2 \sigma T}{\kappa} \tag{1.1}$$

where T is the temperature. One of the primary challenges associated with thermoelectric materials is the competing electrical and thermal properties responsible for high ZT. The thermal conductivity κ may be broken down into κ_{elec}

and κ_{ph} corresponding to electrical and vibrational contributions to the thermal conductivity. Here, κ_{ph} is manifested through phonons, or quasi-particles of vibrational motion, as the primary heat carriers in semi-conductors and insulators. As seen in Equation 1.1, because σ and κ_{elec} scale with electron mobility, κ_{ph} is required to remain low to maintain optimal ZT performance. Consequentially, finding new materials that satisfy the fine-tuned properties of devices such as TE generators are difficult purely by trial-and-error experiments, in-turn slowing the discovery process significantly [3].

In the pursuit of materials discovery for thermal management applications, advancements in computing technology over the past few decades, in tandem with solidified physics and chemistry groundwork of the past century, atomic simulation techniques are in a position to address high-throughput search of materials for energy applications. This is reflected by the exponential increase in research publications involving condensed matter simulations, whereby many complex solids and multi-component liquids are analyzed at the electronic and atomic levels for fundamental understanding of micro- and macroscopic processes [4]. Today, the two conventional atomic simulation methods are molecular dynamics (MD) and density functional theory (DFT), where the former represents atoms as Newtonian systems with parametric atomic interactions and the latter models nuclei and electrons as a quantum mechanical system. Both of these methods have proven invaluable tools for analyzing a plethora of materials for energy applications, including but not limited to superionic lithium ion

conductors [5], molten salts as reactor fuel [6], graphene membranes for natural gas processing [7], and two-dimensional GaN for thermal dissipation [8].

In detail, MD simulations explicitly treat pairwise and many-bodied interactions through analytical potential energy surfaces, which are typically fitted with experiment and/or quantum mechanical simulations such as DFT [9]. The forms of these analytical potentials are known *a priori* based on fundamental chemistry knowledge, commonly involving repulsive, dispersive, van der Waals, and Coulombic interactions depending on the system modeled [10]. Because of their parametric nature, atomic energies and forces are computed based on atomic positions (and bond angles, if required) with high efficiency, allowing simulations of million-scale atomic systems on picosecond to microsecond timescales whereby many important physical and chemical phenomena manifest.

Whereas MD tracks the movements of atoms explicitly, DFT further breaks down systems into nuclei and electrons as a quantum mechanical system. Here, nuclei are much heavier than electrons and are thus treated as ‘frozen’ objects in space, assigning electrons as the only degree of freedom in the system. Formally, this is called the Born-Oppenheimer approximation, which serves to narrow down the total energy or Hamiltonian to the sum of electron-electron interactions, electron-nuclei interactions, and the kinetic energy of electrons [11]. In short, the electron density of the system is self-consistently evaluated for many-bodied systems, yielding the ground state energy *via* the Schrodinger equation. Because electron densities are solved, DFT can provide realistic

representations of bond chemistries and reactions in condensed matter systems, such as atoms, molecules, and periodic systems.

Given MD and DFT as quintessential platforms for atomic-level materials modeling, they also have their share of limitations. Specifically, due to the fixed functional form and innate inflexibility of MD potentials, it is difficult to simultaneously predict multiple properties with decent agreement to experiment. Additionally, MD potentials are material-specific, i.e. they must be refitted with new data for previously unseen materials. Although DFT can model systems of arbitrary elemental compositions, due to the exponential computational cost associated with electronic degrees of freedom, simulating relatively large systems ($> 10^2$ atoms) becomes significantly demanding [12]. Certainly, DFT can handle up to 10^3 atoms with sufficient computational resources, but the exponentially higher costs associated with such systems cannot justify its use for rapid materials screening. Thus, bridging the gap between the two paradigms remains a major subject of research.

A primary route that has taken researchers by storm is machine learning. Due to their demonstrated ability to fit complex non-linear, many dimensional functions at orders of magnitude faster than the traditional enumeration schemes, machine learning methods have been developed to combine the best of MD and DFT methods, i.e. by balancing speed and accuracy. Namely, machine learning potentials (MLPs) capture the explicit electronic-level features from DFT by implicit representation of the potential energy surfaces as functions of the atomic nuclei positions [12]. Due to the purely mathematical nature of machine learning

models, to capture the appropriate physics the accuracy of MLPs strongly depends on the description of the atomic environment surrounding central atoms. It is well known that cartesian coordinates are not well-suited for description of atomic environments, since a) they are not rotationally invariant and cannot physically represent the potential energy, b) the order in which the cartesian coordinates are fed into MLPs is ill-defined and could effectively “confuse” models, c) atoms may leave/enter the designated cutoff range which is incompatible with MLPs containing fixed inputs [13]. As such, several MLPs have been proposed over the past decade designed with various atomic descriptors embedded in a multitude of machine learning algorithms.

Many studies have shown excellent representation of DFT-level energetics and realistic property prediction with MLPs such as the High Dimensional Neural Network Potential (HDNNP) Deep Potential Molecular Dynamics (DeePMD) for a wide array of materials, including but not limited to water, molten salt, metals, semiconductors, organic molecules, superionic conductors, and amorphous structures [14–21]. In the context of phonon property prediction, several studies have been published using MLPs as the force calculator. For instance, Marques *et. al.* reproduced the DFT phonon dispersions for cubic Si and Ge, and Minamitani *et. al.* captured both Si and GaN dispersions and LTC within 5.4% agreement to DFT [22,23]. Typically, all of these studies own a root mean square error (RMSE) of the force predictions from the MLP within 10-100 meV/Å which is similar to order of error from DFT calculations with approximately $\sim 10^3$ faster evaluation time.

Undoubtedly, the robustness of the MLP has the potential to mitigate the current speed-related bottlenecks in the DFT workflow for property prediction. However, to-date a majority of studies with MLPs also share a common denominator in that the models are limited to a material-to-material basis. This is in-part to the interpolative nature of machine learning models such as neural networks, where structures containing positional and species-dependent atomic configurations outside the range of the training set may confuse networks and provide nonsensical outputs [12,24]. Even for crystals sharing similar structure, a change in atomic species may change the density of the structure, shifting the atomic positions. This equivalently pushes the input descriptor for MLPs beyond its narrow knowledge resulting in poor representation of DFT energies and forces. Additionally, many descriptors in current machine learning models suffer from exponential scaling of complexity with respect to the number of elements included in the training set. This creates a significant slowdown in both training speed and accuracy if the model has to represent even a few tens of elements. Furthermore, for training on a variety of similarly structured materials, the question of how many reference *ab initio* calculations are required from each unique material to achieve sufficient accuracy is of concern.

The work presented here serves as a review for the construction and deployment of machine learning-based algorithms for representation of atomic interactions. Primarily, an in-house force field model dubbed the Spatial Density Neural Network Force Field (SDNNFF) is developed for the timely calculation of atomic forces with DFT accuracy. Chapter 2 details the fundamental background

of the SDNNFF, including the development of atomic descriptors, network parameters, and the training dataset. Additionally, Chapter 2 verifies network performance by force evaluation of bulk silicon, diamond, silicon carbide (SiC), and boron arsenide (BAs), whereby phonon properties including dispersions and lattice thermal conductivity are predicted and compared with both DFT and experiment. Chapter 3 shifts to the molecular modeling of eutectic LiF (50% Li) and FLiBe (66% LiF – 33% BeF_2) molten salts with DeePMD potentials for molten salt reactor (MSR) design. Due to the complex formation of local structures in liquid media, accurate representation of molten salt potential energy surfaces is essential to capture the appropriate densities and corresponding dynamical properties, including thermal conductivity and viscosity. Thus, Chapter 3 highlights the importance of training dataset tuning to capture appropriate physics in atomic systems, including the sensitivity of ionic liquids to dispersion interactions. Chapter 4 continues the development of the SDNNFF, renamed to Elemental-SDNNFF, by inclusion of atomic species as part of the input descriptor, allowing for single models to represent materials with constituents spanning the periodic table. The major result here is the modeling of full, half, and quaternary Heusler structures across 55 elements, allowing for rapid prediction of thermodynamic stability and lattice thermal conductivity for > 10^5 unique materials. Results are verified with DFT and experimental values, showcasing the reliability of the model. Additionally, to generate training datasets amidst the expansive space of theoretical materials, active learning schemes are presented as a human-free methodology for model self-improvement. Lastly, Chapter 5

expands on the previous chapters by prediction of ~80,000 cubic materials spanning 63 elements and 16 structural prototypes. Because the Elemental-SDNNFF is a bottoms-up approach, we also study many unique physical properties of crystals, including but not limited to the hydrodynamic phonon transport and lattice thermal conductivity relations to thermal mean square displacements. Overall, this work aims to inform readers a detailed scope of the atomic modeling with machine learning potentials for the purpose of rapid property prediction and discovery of novel materials for future energy applications.

Chapter 2 Spatial Density Neural Network Force Fields with First-Principles Level Accuracy and Application to Thermal Transport[1]

[1] Rodriguez, A., Liu, Y., and Hu, M., 2020, "Spatial Density Neural Network Force Fields with First-Principles Level Accuracy and Application to Thermal Transport," 102(3), p. 35203, Submitted to Physical Review B, 02/11/2020.

Abstract

Constrained by the fixed mathematical form of most empirical potentials used in classical molecular dynamics (MD) simulations, many properties of materials cannot be captured within experimental accuracy. On the other hand, accurate electronic structure calculations based on quantum theory, most notably density functional theory (DFT), are limited to several hundred atoms within picosecond, which makes the method inadequate for modeling systems beyond the nano-scale. A combination of speed from classical MD and fidelity from DFT can be achieved through novel machine learning methods. Herewith, we developed a new approach named spatial density neural network force fields (SDNNFFs) by training neural networks to “learn” and predict DFT-level forces. Our model focuses on the usage of a 3D mesh of density functions, which together act as a mapping of the atomic environment and provides a physical representation of the forces acting on the central atom. Several notable advantages arise from the SDNNFF, including (1) the avoidance of the chain-rule on the total energy and other variables by direct calculation of the forces from the neural network, (2) the ever large $N \times t$ scaling of the training data, where N is the number of atoms in a supercell and t is the number of evaluated structures by first-principles, and (3) the significant reduction in parameters and human effort needed to successfully train a force- and/or property-converged neural network force field. Overall, we focus on modeling DFT-level forces with minimal computational cost and parameterization for rapid prediction of phonon-based properties and future molecular dynamics of large-scale systems. To demonstrate the SDNNFF, we

trained several models on diamond-structures, including bulk silicon (Si), diamond, silicon carbide (SiC), and boron arsenide (BAs), and predicted their phonon dispersions and lattice thermal conductivities using the direct solution to the phonon Boltzmann transport equation (BTE). For phonon properties, we utilized a fitting method for obtaining the second and third order force constants, which outperforms the highly force-sensitive finite displacement method when employing neural network force fields. In comparison to DFT lattice thermal conductivity, we obtained results with unprecedently high precision from our SDNNFF within 0.7% for Si, 6.2% for diamond, 2.76% for SiC, and 7.46% for BAs, with further agreement with experiments. The phonon dispersions from the SDNNFF also matched those from direct DFT and experiments. The developed approach for accurately predicting phonon transport properties of crystalline materials would largely benefit the design of advanced materials with improved performance, such as complex thermoelectric devices and low thermal resistance interfaces for nanoelectronics. Future applications of our SDNNFF model could be extended toward including atomic energy into the algorithm and simulating large-scale heterogeneous systems for quasi-electronic representations for various properties.

2.1 Introduction

The representation of atomic interactions, or the potential energy surface (PES), is key for capturing realistic behavior of materials and their properties. Classical molecular dynamics (MD) and density functional theory (DFT) are two examples of conventional atomistic simulation methods. In MD, the movements of point-particles representing atomic nuclei are dependent on the pairwise or many-bodied potential functions with fixed functional forms. Since interactions vary between atomic species, these potential functions typically consist of adjustable coefficients based on empirical data, scientific knowledge, force-matching with *ab initio* [9], or a combination of the three. An exemplary empirical potential accounts for imaginary electrons surrounding the nuclei and any realistic pair-wise or many-bodied interactions associated with them, e.g. bond stretching, van der Waals forces, torsions, etc. [25]. While the simple functional forms of these potentials allow quick evaluation of the atomic forces and resulting Newtonian dynamics, they are strongly limited in modeling chemical reactions and in predicting multiple properties comparable to experimental data [26]. In addition, any improvement to classical potentials is not straightforward because the dependencies between parameters are not known *a priori* [27]. On the other hand, DFT takes into explicit account the electron densities surrounding nuclei and computes stable structures via iterative minimization of global energy according to the Schrödinger equation [4]. The PES in DFT is represented by the Born-Oppenheimer approximation, where the kinetic energies of nuclei are separated from the electronic Hamiltonian due to their massive sizes in

comparison to electrons [28]. Since DFT is a “bottoms-up” or first-principles approach, an in-depth understanding of material behavior is possible, such as the phonon branch contributions to inverse temperature thermal conductivity for monolayer gallium nitride [8]. However, contrary to the almost linear scaling with the number of atoms in MD, the explicit consideration of electronic degrees of freedom limits DFT to several hundred atoms, which makes the modeling of behaviors beyond the nanoscale, such as the growth of a silicon anode in charging a lithium-ion battery [29], an arduous or otherwise impossible task.

Considering the drawbacks of classical and quantum-based methods, it becomes clear why researchers are pursuing methods that combines the best in both simulation methods by bridging the subatomic and atomic scales together. Recent popularity in machine learning (ML) has inspired researchers to merge this family of computational techniques with atomistic simulation [3]. Inspired by biological processes, ML is oftentimes correlated with the “trial-and-error” mindset in which the predictive power of a model is improved from reference data with little to no human intervention. Here, the reference data “teaches” or fits the model iteratively for the most optimum function between provided inputs and outputs with a high degree of flexibility [30]. Recent computing advances has given strong leeway for ML into materials research, where the generation of data such as the ~10^5 compounds on the Materials Project Genome [31] has led to some major accomplishments [32]. These efforts include, but are not limited to, the discovery of 8 new solid superionic lithium ion conductors [5] and the

prediction of interfacial thermal resistance between two materials related to temperature, coupling strength, and in-plane tensile strains [33].

This paper focuses on the application of ML for the quantitative development of interatomic interactions from accurate quantum-based data. The machine learned potential (MLP) has been a topic of consistent research for more than a decade, with the interest of developing an accurate mapping of the PES from high fidelity data with minimal prior knowledge of the interatomic interactions. Out of the many ML algorithms, feed-forward neural networks were chosen in this study as the algorithm of choice due to their recent popularity in commercial and academic research, and are applied to the atomic simulation field as neural network potentials (NNPs) [34]. These NNPs have yielded agreeable DFT energies and forces for a wide variety of systems, including but certainly not limited to condensed-phase water [35], Li-Si alloys [36], $C_{10}H_{20}$ isomers [19], copper surfaces [17], and amorphous Li_3PO_4 [20]. Other than the composition, structure, and size of modeled systems, papers with NNPs are further broken down by the treatment methods on network inputs prior to training. Since ML algorithms are purely mathematical models with no physical basis, the ability to learn meaningful relationships in data is largely dependent on the quality and consistency of the dataset itself. For NNPs, the representations of the local chemical environment in the training data are critical for accurate predictions of central-atom energies and forces. As an example, suppose the inputs to a NNP are all the cartesian coordinates $\boldsymbol{r}_i = (x_i, y_i, z_i)$ of atoms i, with the total system energy as the output. A rotation of these atoms about an axis should

realistically yield same energy prior to rotation, since the relative distances and angles of the atoms are consistent. Nonetheless, since the values of the atomic coordinates suffer dramatic changes from the rotation, an NNP not trained to support the rotated coordinates will most likely output an energy largely different than that prior to the spatial transformation.

Contrary to most referenced NNPs, including the high-dimensional neural network potential (HDNNP) proposed by Behler et al. [37], here we develop a neural network named the spatial density neural network force field (SDNNFF) that *directly computes and predicts force fields rather than the PES*. Originally as an inspiration to our work, here we discuss the HDNNP briefly. In the HDNNP, the cartesian coordinates of local atoms within a cutoff are transformed into a series of radial and angular symmetry functions that maintains invariance between similar atomic environments, resulting in consistent total energies. The net force $\boldsymbol{F}_i$ acting on atom i is subsequently computed from total energy (E) with respect to its coordinates according to $\boldsymbol{F}_i = -\nabla_i E$. Here, our model diverges from this approach as we pursue the per-atom forces without the energies, since the forces are of interest to compute the interatomic force constants (IFCs) and the dynamics of atomic systems. Direct evaluation of the forces could save computational resources in MD via removal of the total energy and its derivative. In the HDNNP, the per-atom forces are computed by the chain rule and require the evaluation of both the symmetry functions and their derivatives, requiring more CPU time than classical potentials [38]. However, to output the central-atom forces directly, radial and angular symmetry functions as inputs are not a

feasible option since force vectors are indistinguishable in a polar mapping of the environment, e.g. a rotation of an atomic environment would yield an unchanging set of symmetry function values while the force vector components would experience dramatic changes. As such, we decided to *develop our own functional mapping of nearby atoms in 3D space rather than polar space*. Briefly, this is done by generating a mesh of points about every central atom, where predefined density functions measure the local density of the atomic neighbors at each mesh point. Then, the collective group of atomic densities distributed in 3D space serve as representative atomic environments for the neural network to recognize DFT-level force components in-response to relative neighbor positions. Overall, the goal of the SDNNFF is to output first-principles based atomic forces at network speeds by proper treatment of instantaneous atomic coordinates.

The structure of this chapter is organized as follows: A technical background for the SDNNFF and other details for network training are given in sections 2.2-2.5. Then, in section 2.6, the reliability of our SDNNFF model is shown whereby we compute the harmonic (2^{nd} order) and anharmonic (3^{rd} order) force constants and phonon properties (i.e. phonon dispersion and lattice thermal conductivity) of Si from the developed neural networks by conventional finite displacement method (FDM) [39] and ALM [40]. Afterwards, in section 2.7, the ALM method is applied to neural networks trained for similar bulk crystalline materials, including diamond, SiC, and BAs. The resulting phonon properties for each material were evaluated by running convergence tests on various SDNNFF parameters and comparing the phonon properties to those evaluated from DFT-

FDM. In section 2.8, we analyze the results and comment on the SDNNFF considering other recently developed methods. Finally, in section 2.9, we provide a summary and some perspective for future study with the SDNNFF method.

2.2 Theoretical background of the Spatial Density Neural Network Force Field

As previously mentioned, a successful neural network force field must be capable of predicting DFT-level force components as a function of neighboring atomic positions. In a 3D orthogonal xyz coordinate system, because the force vector $\boldsymbol{F}_i = (F_i^x, F_i^y, F_i^z)$ is projected along the reference global axes, the input descriptors should also quantify the atomic positions in the same xyz space. A simple way to do this is to take the cartesian coordinates of nearby atoms and directly insert them into a neural network. However, many issues arise from this approach. Firstly, during neural network training, stability may be compromised from training data that varies greatly due to translated or rotated coordinates. Here, we intend to train networks on force components that differ between spatially transformed varieties of the same system, which is unlike the case when training for the energy. The question then becomes how should the atomic coordinates be fed into the neural network, such that the inputs are consistent to avert large changes in the weights during training? An example of instability from the inputs is the order in which the atomic coordinates are fed into the neural network. Since the weights of a given network are fixed numerical values, inserting two equivalent atomic environments with random ordering for the neighbor coordinates will certainly change the network output. This problem is

trivial for ordered systems with no significant diffusion of atoms, since the network can be fed inputs according to positionally equivalent atoms in space (e.g. by ordering the *xyz* coordinates from greatest to least for all atomic environments as the input). On the other hand, the training of atomic coordinates in disordered systems is infeasible since there are rarely any equivalent atomic structures, which adds an element of randomness to the inputs and hinders the fine tuning of the weights during training. The second point against cartesian coordinates as inputs arises from the entering and exiting of atoms with respect to a certain cutoff radius, which occurs frequently during MD simulations. Treating the cartesian coordinates as inputs is impractical since neural networks are assigned a fixed number of inputs at the start of training, which is incompatible with the variable number of neighbors within the cutoff that may arise during simulations. Again, while this may be negligible for materials with a fixed number of neighbors (e.g. low-temperature solids), the method becomes unreliable for systems where atomic neighbors travel far from the central atom. For the compatibility of neural network force fields with any arbitrary system, the focus is to represent the space surrounding the central atom independent of any periodicity or possible atomic environments found in DFT or MD.

Inspired by other works on atomic representations, we decided to apply the concept of symmetry functions to represent 3D variations in the environment to avoid the issues associated with cartesian coordinates. As seen in Figure 2.1, First, we gather the crystal structure and its per-atom forces from first-principles, then for each atom we describe the local atomic environments by density

functions with a finite neighbor cutoff radius. The result is a atomic density mapping of the environment which can be fed into a neural network. Specifically, we train the network on density values found at each grid point to obtain the force vector for various atomic configurations. Then, we can make DFT-level force predictions at significantly improved speeds for applications such as rapid phonon property predictions done here (shown: predicted phonon dispersion and lattice thermal conductivity of Si) or in the future simulate MD at DFT accuracy (shown: lithium ion diffusion through silicon anode [29]).

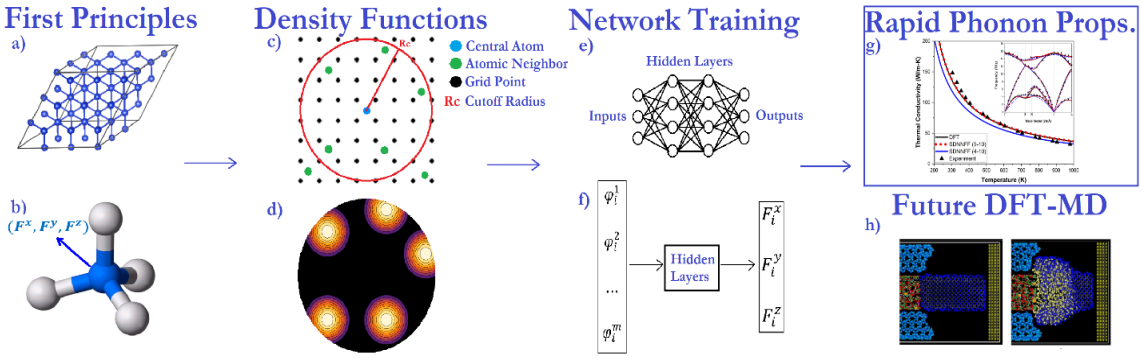

Figure 2.1 Workflow for the Spatial Density Neural Network Force Field.

We follow the simple forms presented by Huang *et. al.* [41], where equally spaced cosine functions serve as mathematical indicators for the atomic occupation surrounding a central atom. This form is especially desirable considering that only three adjustable hyperparameters are present, the cutoff radius R_c, the grid resolution k, and the density function width D. The first two parameters are physically intuitive since R_c determines the distance truncation of the force contribution to the central atom and k provides the spatial resolution of

the grid points surrounding the central atom. The third parameter D measures the range of each cosine function, and we show later a universal value for given structures can be found using network convergence testing. Therefore, a good representation of the local environment and the resulting central-atom forces depends on the convergence of the force error and material properties by adjustment of R_c and k. Additionally, the grid-centered cosine functions were modified to transform *xyz* coordinates to a form independent of the quantity or periodicity of the neighboring atoms. In other words, our method was designed to represent any given atomic environment, provided that the inputs to the neural network are treated as fixed observers of the neighbors, and the quantity of atoms inside R_c does not change the number of inputs to the neural network. Therefore, a high flexibility is achieved in these SDNNFF functions, formally named density functions, which allow neural networks to capture the complex force fields involved in a wide variety of systems.

The theoretical background for the SDNNFF density functions is provided here. For a 3D grid with $k > 1$ grid divisions, a given spherical cutoff radius R_c, $\boldsymbol{R}_n = (x_n, y_n, z_n)$ as the coordinates of the n^{th} atom relative to a given central atom, and $\boldsymbol{\alpha} = (\alpha^x, \alpha^y, \alpha^z)$ as the grid positions with each component ranging from 1 to k, the proposed spherical density function is shown in Equation 2.4, with Equations 2.1, 2.2, and 2.3 provided as definitions

$$L = \frac{2R_c}{k\text{-}1} \tag{2.1}$$

$$\boldsymbol{R}_g = -R_c + (\boldsymbol{\alpha} - 1) \times L \tag{2.2}$$

$$R_n^{\alpha} = \sqrt{\left(x_n - R_g^x\right)^2 + \left(y_n - R_g^y\right)^2 + \left(z_n - R_g^z\right)^2} \tag{2.3}$$

$$If\ |\boldsymbol{R}_g| = \sqrt{R_g^{x\,2} + R_g^{y\,2} + R_g^{z\,2}} \le R_c$$

$$\varphi_{\alpha} = \begin{cases} \sum_{n=1}^{N_r} \frac{1}{2} * \left(\cos\left(\frac{\pi}{D \times L} R_n^{\alpha}\right) + 1\right) & if\ R_n^{\alpha} < D \times L \\ 0\ otherwise \end{cases} \tag{2.4}$$

where

L the distance between equally spaced grid points

$\boldsymbol{R}_g$ the cartesian coordinates of grid point $\boldsymbol{\alpha}$ with $\boldsymbol{R}_g = (R_g^x, R_g^y, R_g^z)$

R_n^{α} the distance between an atom n and a grid point $\boldsymbol{\alpha}$

N_r number of atoms within R_c

φ_{α} the local atomic density value at grid point $\boldsymbol{\alpha}$

The basic requirement of the SDNNFF density functions is to provide a description of the 3D atomic occupation in space. Hence, the grid space surrounding the central atom is defined by the diameter of the cutoff sphere, or $2R_c$. This range is then divided into k partitions, generating the inter-grid distance L according to Equation 2.1. In Equation 2.2, the real-space coordinates of the grid points $\boldsymbol{R}_g$ are given from $-R_c$ at $\alpha^x, \alpha^y, \alpha^z = 1$ to R_c at $\alpha^x, \alpha^y, \alpha^z = k$ in each of the *x-/y-/z-* directions. Equation 2.3 provides R_n^{α} as the relative distance between a neighboring atom n and a given grid point $\boldsymbol{\alpha}$, which measures the spherical vicinity of atoms with respect to a grid point.

These definitions lead to the density function in Equation 2.4 provided as a cosine function ranging from zero to one varying with radial distance R_n^{α} , as seen by the 2D representation in Figure 2.2. In the figure, the function has a radius $D \times L = 1$ centered at $\boldsymbol{R}_g = (0, 0)$. Atoms in the black region would result in a value of zero, increasing to unity as the atom approaches $\boldsymbol{R}_g$.

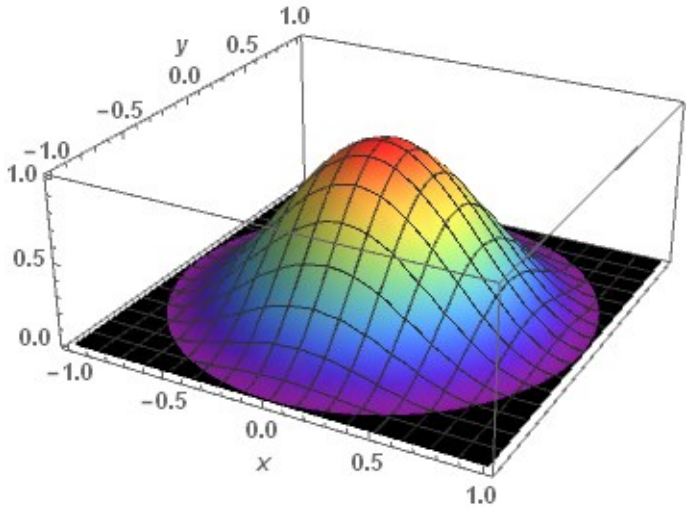

Figure 2.2 The shape of the per-node density function in 2D.

The density function φ_{α} considers the radial influence of up to N_r atoms inside R_c, and their contributions are summed into one value per grid point to provide a quasi-density of atoms in space. The summation also acts as a necessary conditioning step for usage in neural networks since the number of atoms within the spherical cutoff becomes independent of the number of inputs to the network. Here, the number of α translates to an equal number of fixed inputs to the

SDNNFF and scales with the grid resolution k. Additionally, the “if” statement preceding the density function in Equation 2.4 is a physical treatment, where only the grid points within the atomic cutoff are considered due to two primary reasons: 1) Like in MD, the forces are strictly dependent on the radial proximity between the central atom and the neighbors, which is represented by a spherical region as opposed to a cubic one. Instead, we limit the grids to those within R_c, permitting a physical description of the forces in the neural network. 2) The number of grids considered is significantly reduced when truncating to the grids within the sphere of radius R_c versus a cube with side length $2R_c$. This effectively reduces the number of necessary inputs to the neural network by a significant margin (at least by volume fraction $\frac{\frac{4}{3}\pi R_c^3}{(2R_c)^3} \times 100\% = 52.36\%$). In-turn, we can increase the grid resolution k to describe the atomic environment more precisely within the cutoff sphere.

A discussion of the density function parameter D is necessary to understand its role in the SDNNFF. As per the definition in Equation 2.4, the density function ranges from unity at its center to zero at a distance $R_n^{\alpha} = D \times L$. Beyond this, the function is treated as zero throughout the rest of the domain, providing a smooth curve delineating the local concentration of atoms at each grid point. Here, the purpose of the cutoff factor D is to control the size of the local spherical domain that scales with a constant grid length L. The choice in D is important to guarantee no gaps in the domain between adjacent density functions, where a small value may cause the density functions to completely neglect the presence of certain neighbors critical for computing the forces. On

the other hand, a value that is excessively large desensitizes the φ_α set to smaller atomic displacements due to the greater width of the local density functions. Furthermore, because all the density functions here are uniform in size and shape, *D* controls the degree of overlap with nearby density functions and in-turn serves as an implicit representation of the angular positions of atoms. While the density functions only depend on radial distance from their center and not on angular values, multiple sufficiently wide density functions provide a good interpolation of the atomic densities between the points in the local environment. The idea is analogous to GPS systems, where multiple orbiting satellites transmit long-range signals to triangulate a position on the surface of the Earth. Here, the value φ_α from each grid point "satellite" is simultaneously fed into the neural network, and as a multi-dimensional non-linear model, we expect the network to recognize the set of φ_α values positioned at equally spaced 3D points as a unique description of the atomic environment as seen above in Figure 2.1. As such, in a later section 2.5, we find the optimal value for *D* by comparing the performance of the resulting neural networks trained on DFT forces.

2.3 Data Generation

For training the SDNNFF, data acquisition is vital for the force prediction of all possible atomic configurations found during long-timescale simulations. Realistically, we would require a nearly infinite number of datapoints to cover all possible combinations of atomic environments. However, since neural networks excel at interpolation, the training data is at minimum required to cover the possible range of inputs [13]. This implies that neural networks must be trained to

suit the application, and because the maximum displacement of atoms scales with the temperature of the system, the coverage of possible atomic configurations can be achieved by training neural networks on a variety of temperatures. Here, the goal was to predict the second- and third-order constants of crystalline materials to find phonon-based properties with *ab initio* accuracy. Since the force constants require small atomic displacements less than 0.03 Å, we decided it would be enough to train neural networks on data far from the melting temperatures (i.e. solid crystalline phase).

Six temperatures were arbitrarily chosen, namely 0 K, 50 K, 100 K, 300 K, 500 K, and 800 K, to train the material specific SDNNFFs on a sufficiently wide range of atomic configurations. Here, the 0 K configurations are perfect crystal structures. To model a variety of atomic environments, we use *ab initio* MD (AIMD) with DFT from the Vienna *ab initio* simulation package (VASP) with the projector augmented wave (PAW) method and a 10^{-6} eV energy convergence criterion to generate the structures. The AIMD simulations were sped-up with a low 1×1×1 *k*-space resolution since at this stage there was no concern for the accuracy of the forces but for the rapid generation of the atomic configurations at various temperatures. For each crystal at non-zero temperature, 5000 low-resolution steps were ran for 3×3×3 primitive supercells (54 atoms in total) containing Si, diamond, SiC, and BAs. Also, a timestep of 1 fs was chosen for Si, SiC, and BAs, while a timestep of 0.5 fs was chosen for diamond due to its known high vibrational frequency. Then, 1250 structures were randomly chosen out of the 5000 AIMD configurations and pooled together to avoid biased training,

where high resolution 4×4×4 *k*-space static DFT calculations with 10^{-6} eV energy and 10^{-3} eV/Å force convergence criterions were performed on each structure. A total of 4×10^{5} non-zero temperature atomic configurations for each Si, diamond, SiC, and BAs were obtained. An arbitrary 1250 0 K (perfect crystal) structures were added to the data pool for the inclusion of equilibrium structures during training. Note that the 0 K configuration for each material was ran in static DFT to check for equilibrium (i.e. 0 eV/Å forces) and was not ran in AIMD.

2.4 Neural Network Details

After the static DFT calculations by VASP, post-processing for the output data was performed in Python 3.6.8 with the NumPy library package [42]. Our in-house Python script converts the generated atomic configurations from the VASP output to the SDNNFF input. The neural network inputs were processed according to Equations 2.1 through 2.4 and saved in a separate file along with the corresponding DFT forces. For training the SDNNFF, the TensorFlow library for Python was utilized [43]. To compare with other papers containing NNPs, all neural networks here were similarly constructed with two hidden layers and 500 nodes each with a default mini-batch size of 32 [41], [23]. The SDNNFF data from our earlier Python script was read and split into 80% training data, 10% validation data, and 10% testing data. In addition, the input density function column vector and output force vector were normalized so that they ranged from -1 to 1 according to

$$I_{network} = 2\frac{I - I_{min}}{I_{max} - I_{min}} - 1 \tag{2.5}$$

where I is a density function or force component from DFT and $I_{network}$ is the converted input/output for the neural network [13]. The values for I_{max} and I_{min} were taken as the maximum/minimum values for each component out of the entire dataset. As such, these values were saved in a separate file since all future uses of the network require the reapplication of Equation 2.5 to the inputs and outputs for consistency with the trained weights. The He uniform initializer was used to generate the initial weights and biases of the network [44], the Adam algorithm was used as the optimizer for updating the weights with a learning rate of 0.0001 [45], and the exponential linear unit (ELU) was implemented as the activation function of choice [46]. The custom loss function Γ was defined as an *RMSE* of the three force components predicted by the neural network F^k_{NN} against the DFT force components F^k_{DFT}, or

$$\Gamma = \sqrt{\frac{1}{3}\sum_{k=x,y,z}\left(F^k_{NN} - F^k_{DFT}\right)^2} \tag{2.6}$$

After training, the testing loss was computed across the entire testing set with Equation 2.6 for comparison with other networks trained on various cutoffs and grid resolutions.

2.5 Determination of the Local Density Cutoff Factor

With the data generated and the network parameters defined, we proceeded to investigate the necessary value for the cutoff factor D in the density

functions. At this stage, we decided to start with high-resolution networks $k = 12$ (672 inputs) and a large R_c containing up to the 5th neighbors for Si (6 Å), diamond (3.93 Å), SiC (4.82 Å), and BAs (5.3 Å). Note that these cutoff distances are approximately 0.05 Å greater than the 5th neighbor distance to guarantee the capturing of displaced atoms when later computing the force constants. In addition, the number of inputs is quite large in comparison to recent reports for HDNNP (i.e. 41 inputs) [23]. Later these parameters would be reduced to decrease the number of required density function calculations for the network input and to improve the efficiency of the force calculator, but for now these are simply test cases to minimize noise error in analyzing D. To observe if a universal value exists for D in the case of diamond-like materials, we trained multiple high-resolution networks for each material with different values of D and compared their testing *RMSE* values according to Eq. (6). Specifically, we sampled values $D = \frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{2}, 1, \sqrt{2}, \sqrt{3}$, and 2, which correspond to a $D \times L$ equal to the grid length ($D = L$), the diagonal of the grid length in 2D ($D = \sqrt{2}L$), and the diagonal of the grid length in 3D $D = \sqrt{3}L$ in multiples of $\frac{1}{2}$, unity, and 2.

From Figure 2.3, it is clearly seen that D plays a significant role in the force accuracy of the network. Specifically, we can see that there is a converged *RMSE* value at $D = \sqrt{2}$ for all materials examined here. From an intuitive standpoint, the convergence is due to the overlapping of density functions with those at nearby grid points, which supports the earlier idea that the angular

information is preserved in the set of input density functions and therefore the representation of the atomic environment is improved.

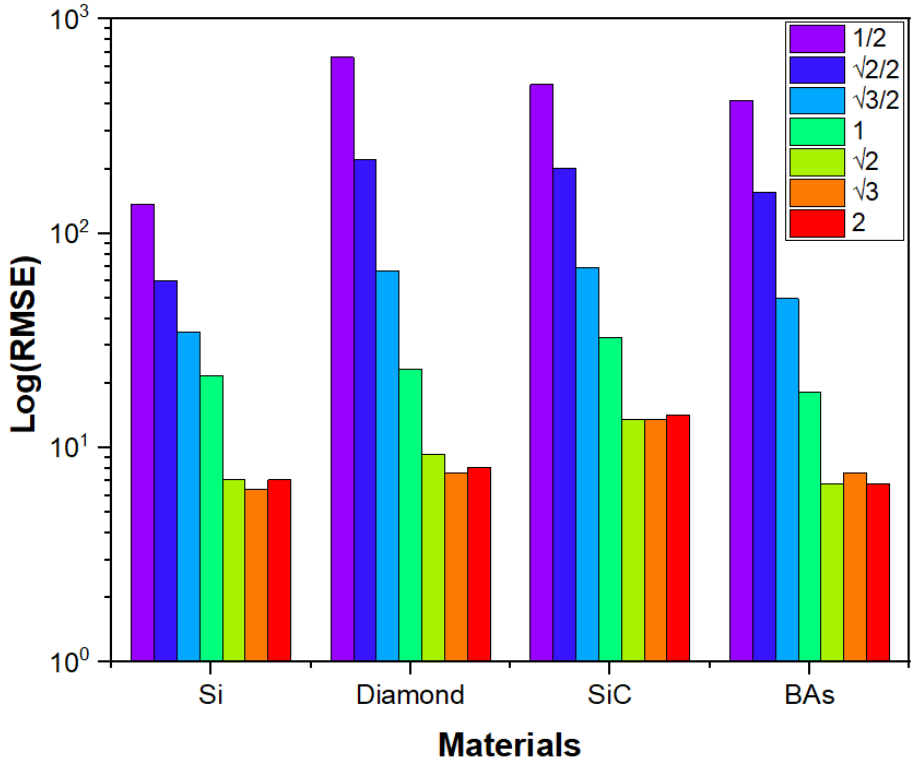

Figure 2.3 The logarithm of the testing RMSE for various cutoff factors for bulk Si.

The overlap begins at D = 1, where the edges of the density functions lie on nearby grid points. The results seem to converge at a slightly higher value of D = $\sqrt{2}$, where the grid points are well within the influence of multiple neighboring density functions. Recognizing the negligible differences in *RMSE* when $D > \sqrt{2}$, we decided to arbitrarily use $D = \sqrt{3}$ for all material neural networks investigated here. These *RMSE* differences could be due to statistical noise in training the network, and thus $D = \sqrt{3}$ was chosen purely to guarantee force convergence

due to the overlap of density functions. However, we recommend in future studies that a convergence test for D should be performed for new structures other than diamond, considering how the packing of atomic neighbors may impose other requirements on the range of the density functions for converged forces. In addition, care should be taken in choosing a large cutoff factor, since the density function values could become “clouded” from the excessive number of atoms included in the summation, thus generating difficulty for the network to detect atomic configurations from the φ_α set. Furthermore, the inclusion of more atoms in the summation for φ_α can increase computational cost, since more cosines are required to be computed for every atom inside $D \times L$ at each grid point.

For the trained networks with $D = \sqrt{3}$, we also observed the learning curves to compare the training and validation losses. Normally we expect the validation loss to decay to a minimum then rise after a given number of epochs (training cycles), which is a symptom for overfitting. During training, the network with the lowest validation error is saved since this model can best generalize untrained configurations. However, we noticed that none of our networks experienced this phenomenon, rather the validation curves were seen to continue to slowly decay with no bound as seen in Figure 2.4. While overfitting could had occurred at an epoch much larger than 500, at this point the two curves were visually close to an asymptote, so we stop at 500 epochs for all networks to save computational resources.

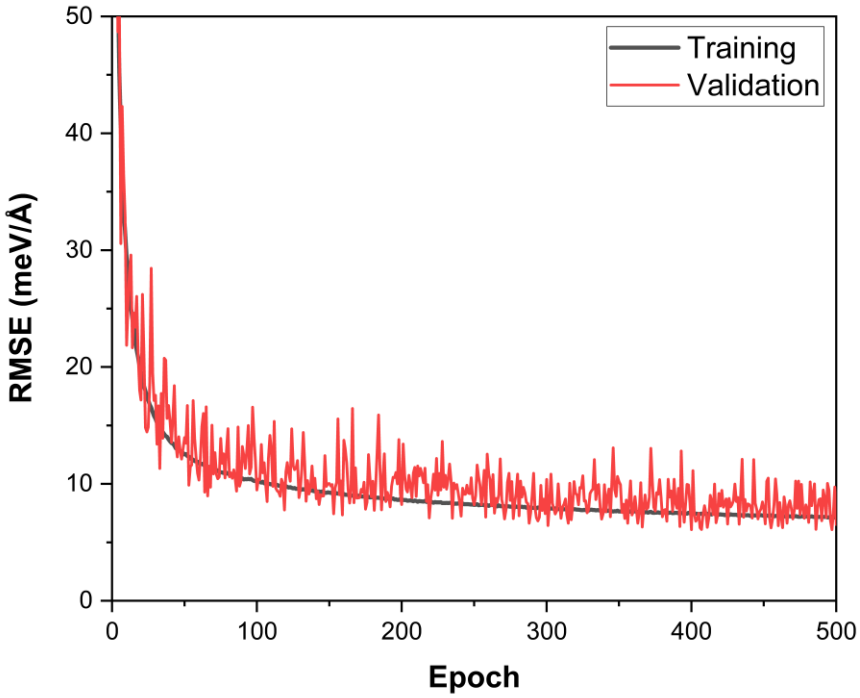

Figure 2.4 The training and validation learning curves for the high-resolution Si network.

2.6 Generation of the Force Constants and Phonon Properties

The performance of the neural networks was primarily determined by their ability to reproduce DFT-level phonon properties. Initially we investigated the properties by calculation of the force constants via conventional FDM. Here, we started with Si as a benchmark case since it is a well-known semiconductor. In this paper, we used the phonopy [47] and phono3py [48] packages for evaluating the 2nd and 3rd order force constants and for obtaining phonon dispersions as well as lattice thermal conductivity. We used the same high-resolution Si network as developed in the previous section to see if our method could yield agreeable phonon properties to those from direct DFT. With FDM, 181 3×3×3 primitive

supercell structures of Si were generated with the default displacement of 0.03 Å. We computed the forces for these structures by static DFT and by the trained high-resolution neural network for cross comparison. Then, the 2nd and 3rd order force constants were computed followed by the phonon dispersion and lattice thermal conductivity. The lattice thermal conductivity was obtained with a 20×20×20 *q*-point mesh using the direct solution of linearized BTE.

As seen from Figures 2.5 and 2.6, while the phonon dispersion from the SDNNFF-FDM agrees relatively well with that from DFT-FDM, the comparison of thermal conductivity between the two force calculators for FDM was significantly worse. We assumed that the phonon dispersions, which only require 2nd order IFCs, are much less sensitive to the force noise error from the neural network than the thermal conductivity, which require both the 2nd and 3rd order IFCs.

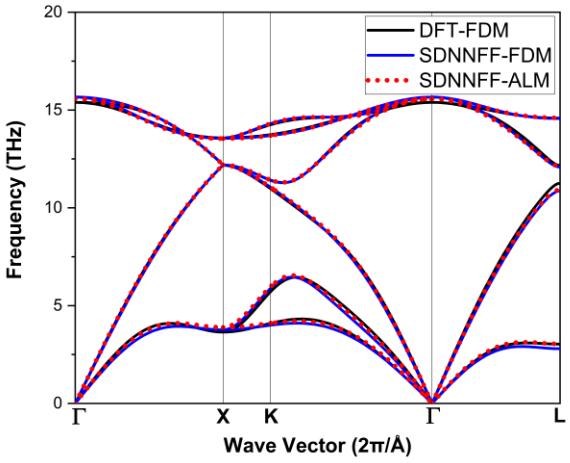

Figure 2.5 The phonon dispersion of bulk Si evaluated by the model and density functional theory.

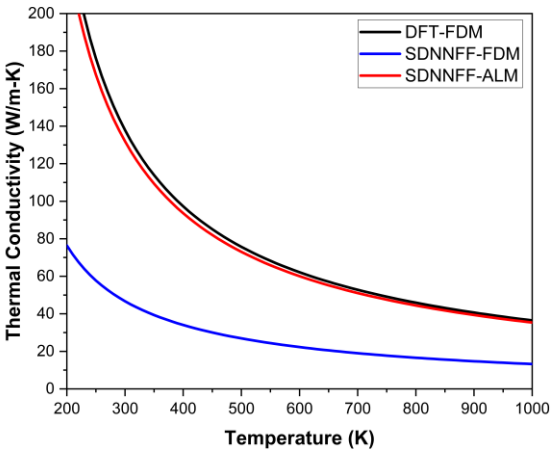

Figure 2.6 The temperature-dependent lattice conductivity of bulk Si by the model and density functional theory.

To prove this, we performed computer experiments by artificially imposing random noise into the DFT forces to see if the phonon properties deteriorate in a similar manner. Noise values of up to $\pm 10^{-n}$ eV/Å, with $n = 2, 3, 4, 5$, were added to the DFT forces. The 2nd and 3rd order force constants, along with the phonon dispersions and conductivities, were obtained by the same procedure as before but with the noise-induced forces. We show the resulting plots in Figures 2.7 and 3.8, where we obtained a converged phonon dispersion with $n = 3$, while convergence for the thermal conductivity occurred at lower noise levels of $n = 5$. These figures agree well with our earlier speculation, where the thermal conductivity requires a demanding 10^{-5} eV/Å or 0.01 meV/Å force accuracy to match DFT results which is well below the capabilities of the neural network even at a high resolution.

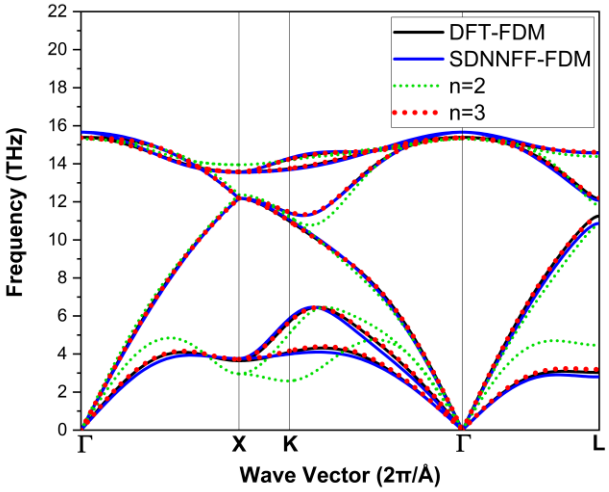

Figure 2.7 The phonon dispersion of bulk Si with the model and two noise-induced dispersions with 10^{-n} noise level from density functional theory.

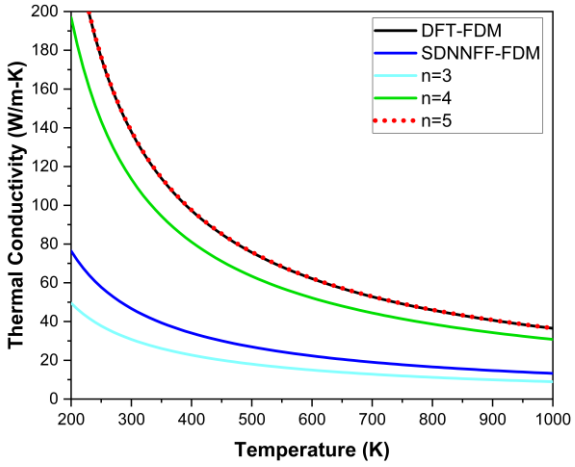

Figure 2.8 The temperature-dependent lattice conductivity of bulk Si with the model and two noise-induced dispersions with 10^{-n} noise level from density functional theory.

In Figure 2.9 we summarize this point by plotting the thermal conductivity of Si at 300 K against the artificial noise in DFT. We observe again that the thermal conductivity from DFT converges at a noise level of $n = 5$ (up to $n = 8$ plotted to show convergence) while the thermal conductivity from the SDNNFF was consistent with DFT noise levels between $n = 3$ and $n = 4$, sharing values far below pure DFT results.

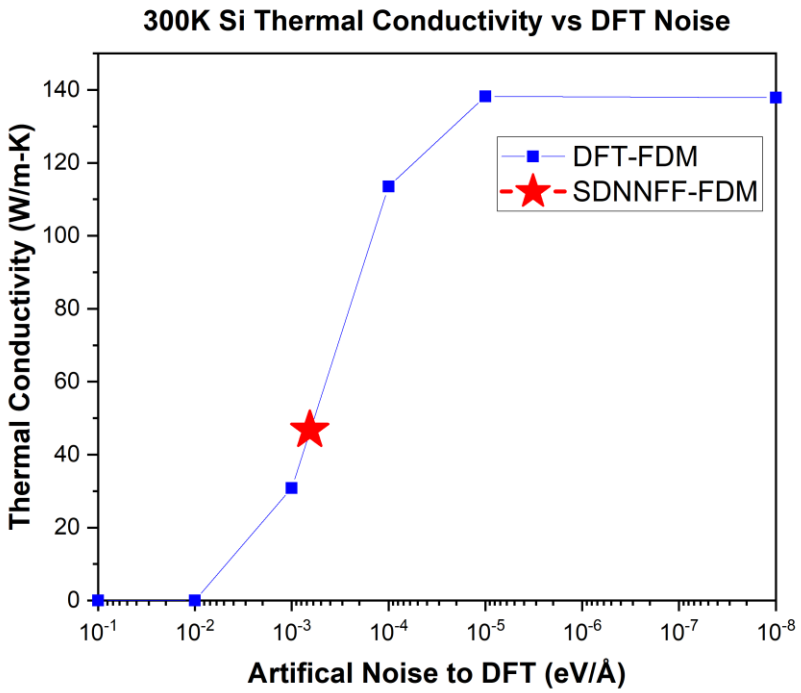

Figure 2.9 300 K thermal conductivity of bulk Si versus artificial noise.

We quickly realized that FDM was not suitable for the evaluation of the force constants via SDNNFF due to the observable strong force sensitivity. While the possibility of reducing the *RMSE* exists by increasing the grid resolution, the number of inputs to the network would have quickly made the method demanding

for force evaluation of several hundred supercells. Instead, we decided to use an alternative force constant calculator, namely the ALM method. While the details of the method can be found from Tadano *et. al.* [40], a short discussion is made here. In FDM, the force constants are computed by systematic displacement patterns in crystalline structures, where a minimum number of unique displacement patterns or structures are generated. From a series expansion on the potential energy, the 2nd order force constants require one atomic displacement, the 3rd order force constants require a pair of atomic displacements, and further order terms require more simultaneous atomic displacements [49]. Then, the electronic structure calculations are done on all resulting structures for the DFT forces and the 2nd and 3rd order force constants can be computed. However, contrary to FDM, ALM approaches the force constants calculation stochastically by fitting the force constants rather than solving for them directly. This is done by applying random displacements to all atoms in the supercell up to a user-specified amplitude. The result is an arbitrary number of randomly displaced structures which, with statistical fitting methods, yields the force constants.

With the ALM extension in phonopy/3py, we were able to evaluate the force constants from the same neural network for Si. We chose to randomly displace atoms by 0.03 Å for an arbitrary 1000 structures. Here, the neural network evaluated the forces for the Si 3×3×3 primitive supercell ~10^3 times faster than DFT, so the force calculation for 1000 structures was not a concern. In addition, we expect the neural network cost to scale linearly with N atoms

since the calculation only depends on the relatively constant number of neighbors in the local environment, similar to MD and contrary to the scaling with the cubic electron count (~N_e^3) in DFT [38]. Viewing the resulting phonon properties as seen in Figures 2.5 and 2.6, we realized the exceptional performance in replicating the DFT-FDM results from the SDNNFF-ALM force constants. The most likely reason for the improvement is in the statistical cancellation of noise provided by ALM, where many atomic displacements (and therefore unique interactions) are necessary to eliminate small discrepancies in the predicted forces. Indeed, FDM is limited to a fixed number of displaced structures with only one to two atoms displaced at a time for the 2nd and 3rd order force constants. The small sample of displacements increases the demand for accurate forces, as we have seen previously with the artificial disturbance of the DFT forces in FDM structures. Additionally, the finite displacements in DFT generates a non-zero force for all the atoms in the unit cell, whereas a real-space force cutoff approach such as the SDNNFF generates equilibrium forces (i.e. 0 ± noise due to model error) for atoms outside the cutoff of the displaced atom(s). As such, the truncation of the forces outside the cutoff is possibly another source of the incompatibility between the SDNNFF and FDM. Note, we strongly wish to emphasize the difference between the real-space force cutoff R_c and the pairwise atomic displacement cutoff used often in FDM code, where the latter is used to reduce the cost of 3rd order IFC calculations by neglecting structures with distances between displaced atoms beyond the specified cutoff. In practice, the pairwise displacement cutoff in DFT-FDM does not define a force cutoff for atoms

in the cell, since the force is dependent on the gradient of the electronic Hamiltonian [11] and is attributed by all atoms in the supercell regardless of the pairwise displacement cutoff. Accounting for both information loss and inherent noise of the SDNNFF, the reduced sensitivity of the force constants led us to prioritize ALM for the evaluation of phonon properties.

2.7 Application to Thermal Transport

With success in reproducing Si DFT results, the next step is to evaluate the phonon properties of the suggested materials by adjustment of the SDNNFF parameters. Similar to bulk Si, we began by testing the high-resolution neural networks for bulk diamond, SiC, and BAs using ALM. We also developed other networks with combinations of 4^{th} and 5^{th} neighbor cutoffs and $k = 8$, 9, 10, and 12 for convergence to DFT-FDM results with a minimum number of grid points and atomic neighbors. Note that we again applied a cutoff that is 0.05 Å greater than that for the 4^{th} and 5^{th} neighbor distances to account for displacements in ALM. In addition, we noticed similar learning curve behaviors as seen in Figure 2.4 in these networks and do not show them here for brevity. However, the final testing *RMSE* values for the various atomic cutoffs and resolutions are displayed in Figure 2.10 for comparison. We also provide the number of input nodes (or the length of the φ_{α} set) corresponding to each k in Table 2.1.

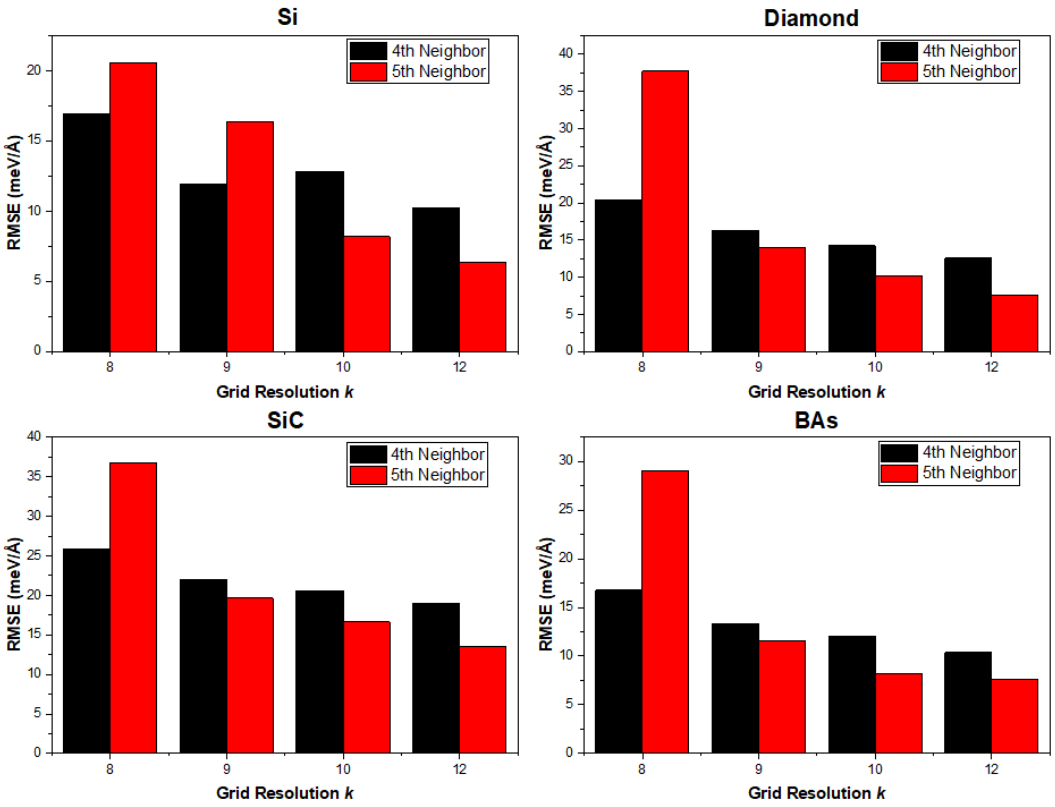

Figure 2.10 The resulting testing *RMSE* values from the material-specific models at various atomic neighbor cutoffs and grid resolutions.

Table 2.1 The number of grids/network inputs to the grid resolution

Grid resolution (*k*)	8	9	10	12
Number of grids / network inputs	160	257	360	672

Because of the previous disagreement with the thermal conductivity via FDM, we prioritized our analysis on the conductivity and show this first in Figure 2.11. Here, we truncated the converged solutions to DFT-FDM results for display purposes, and only show one curve for each neighbor cutoff. Plots that are not shown for a given combination of neighbor cutoff and *k* have either equivalent or

worse agreement to DFT-FDM. Afterwards, the phonon dispersions corresponding to the shown thermal conductivity curves are plotted against DFT-FDM dispersions in Figure 2.12. For the phonon dispersion of SiC, we applied a non-analytical term correction due to its known LO-TO splitting near the gamma point [50]. We also provide experimental datapoints in Figures 2.11 and 2.12 for Si [51,52], diamond [53–56], SiC [57–59], and BAs [60,61] for comparison except for the thermal conductivity of BAs, which requires four-phonon scattering theory to match experiment [62]. All DFT curves were computed from FDM and all SDNNFF curves were computed from ALM unless otherwise specified. We also provide the force line comparison plots for the same SDNNFF networks in Figure 2.13.

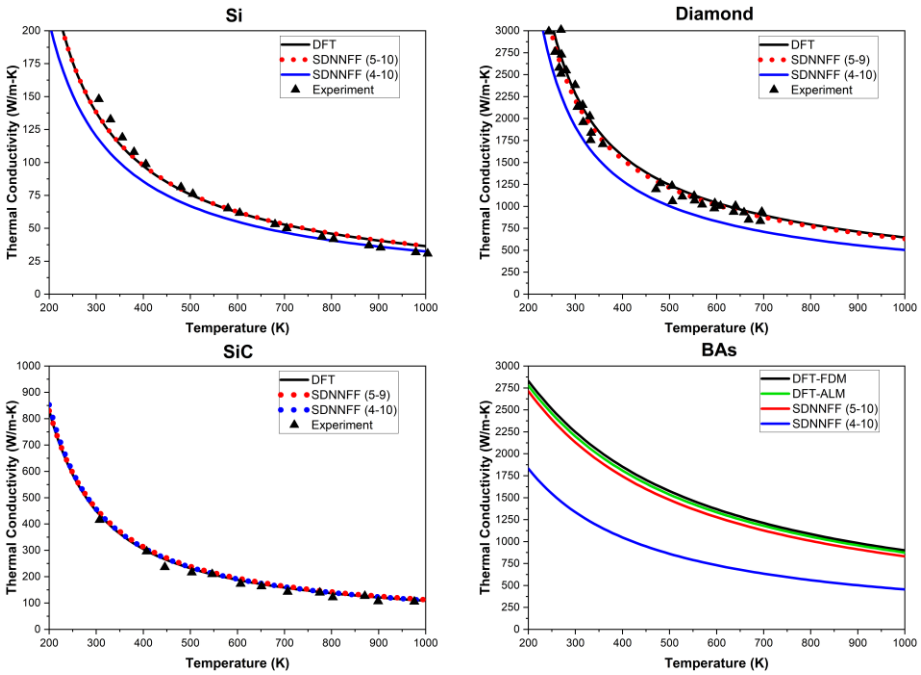

Figure 2.11 The temperature dependent lattice thermal conductivities of all four diamond-structure materials studied here.

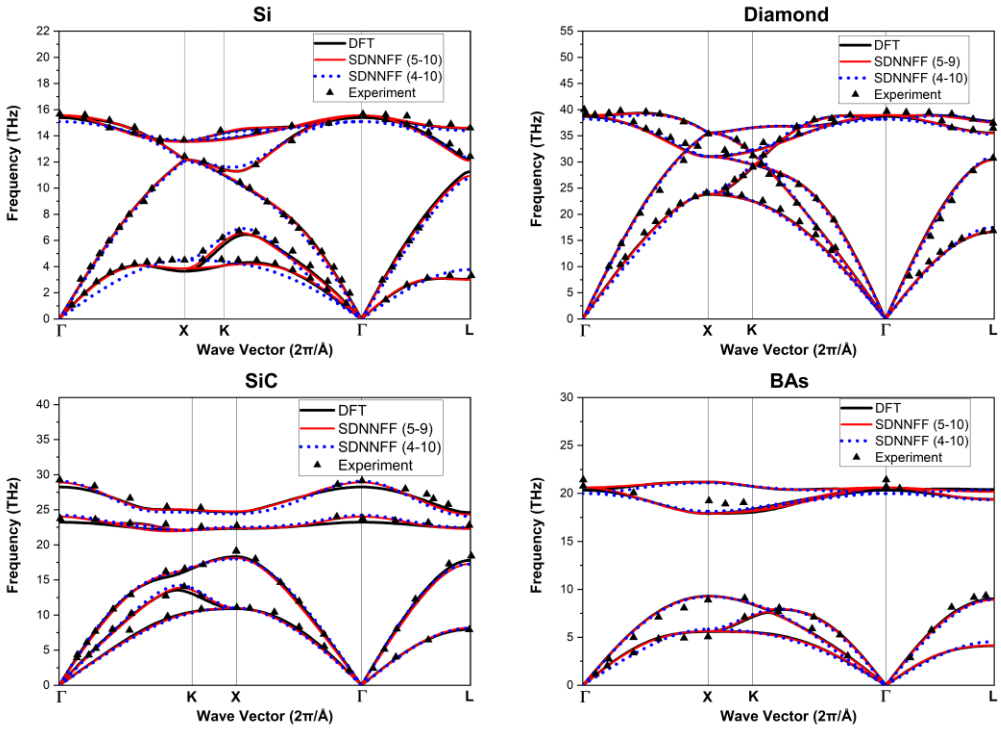

Figure 2.12 The phonon dispersions of all four diamond-structure materials studied here.

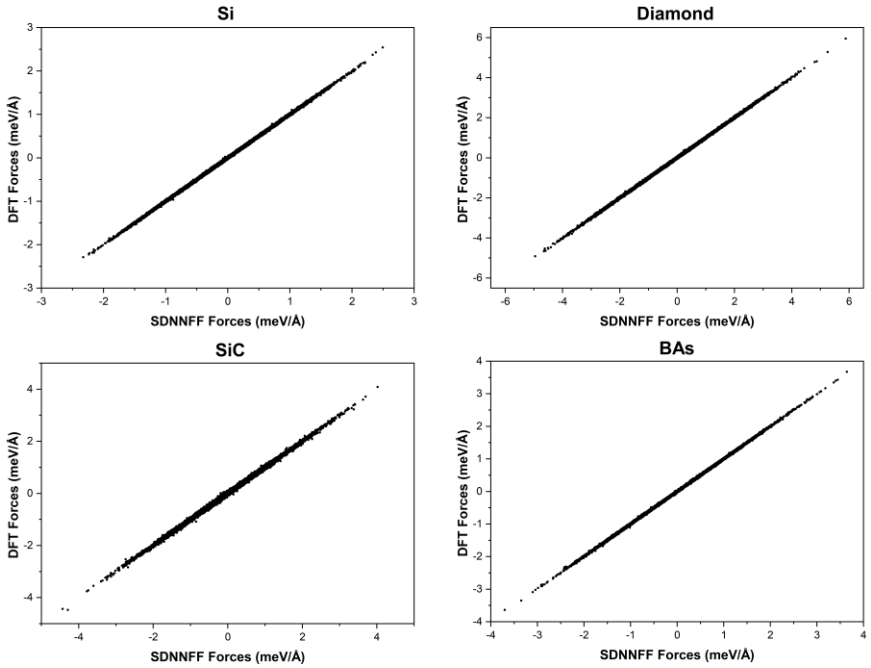

Figure 2.13 The comparison between the target and predicted forces from the testing set.

2.8 Discussion

As seen from Figure 2.11, we obtained excellent agreement in the lattice thermal conductivity from the SDNNFF forces with the ALM IFC calculator. A maximum percent difference of 0.7% for the 5-10 Si network, 6.2% for the 5-9 diamond network, 2.76% for the 4-10 SiC network, and 7.46% for the 5-10 BAs network in comparison to the DFT-FDM curves was achieved. Considering BAs as the highest percent error for the thermal conductivity, we decided to also plot DFT-ALM to see if the percent error originates from the IFC calculator (FDM vs. ALM) rather than the atomic force accuracy (DFT vs. SDNNFF). Here we ran 300 ALM structures for BAs instead of the previous 1000 due to the demanding

number of DFT calculations and the expectation that less structures are required for fitting the force constants with accurate DFT forces. The resulting thermal conductivity from DFT-ALM was slightly lower than that of DFT-FDM and improved the agreement with SDNNFF-ALM to 4.81%. While we do not analyze the effects of the force constant calculators on the thermal conductivity here, the comparison between FDM and ALM shows that the disagreement is not purely due to the SDNNFF noise. Other sources of error that may have contributed to the resulting percent difference in the SDNNFF thermal conductivity include the number of reference DFT structures in the training set, the temperature range of the generated AIMD structures, the neural network training parameters (e.g. number of layers, nodes, activation functions, etc.), the number of random ALM structures, and the amplitude of the random displacement. However, here we focused on providing a generalization of the SDNNFF performance, and as such did not share excessive concern for tuning these parameters for each material. We only focused on tuning the SDNNFF parameters (i.e. the neighbor cutoff, the grid resolution, and the local cutoff factor) to showcase its potential to successfully reproduce the conventional DFT-FDM and experimental results. Given the various known sources of errors, we were satisfied with the less than 10% error in thermal conductivity and the agreement with both DFT-FDM and experimental trends.

From the predicted phonon properties, we see that the combination of the number of neighbors and the grid resolution is not trivial and has serious impact when reproducing DFT results. Here, the resolution represents the number of

inputs to the neural network, i.e. a higher number of grids equals more information for the network to identify the atomic environment and the corresponding central-atom forces. However, we can see that the number of inputs required for the network to sufficiently learn the forces reaches an upper limit, since *k*-values greater than what were provided for each material in Figure 2.11 do not improve the results. In other words, the additional information obtained from higher grid resolutions becomes redundant for predicting the thermal conductivities accurately, even though the network *RMSE* continues to drop as seen in Figure 2.10. On the other hand, the number of neighbors provides physical meaning to the neural network, where atoms beyond the cutoff are neglected in the contributions to the central-atom forces. At minimum, the cutoff should be large enough to where atomic contributions outside the cutoff are negligible or otherwise inconsequential for property prediction. In Figure 2.11, we noticed for the thermal conductivity that the 5^{th} neighbor contribution is quite significant in Si, diamond, and BAs. However, the 4^{th} neighbors seem to have nearly equal contributions to the thermal conductivity in SiC. For the phonon dispersions in Figure 2.12, the 5^{th} neighbor contributions for Si are necessary to match DFT as seen by limited agreement in the 4^{th} neighbors with the low group velocity near the Brillouin zone center. Otherwise, the phonon dispersions truncated at the 4^{th} neighbors for the other materials are relatively close to the DFT-FDM dispersions. Overall, the SDNNFF parameters R_c and k seem non-transferrable between materials and require further convergence testing for optimal network performances.

Summarizing the performance of our method, a trained SDNNFF is quite capable of reproducing DFT-level and experimental phonon properties with significantly lower cost compared to full *ab initio.* For comparison with other papers, the *RMSE* for the Si network containing the minimum neighbor cutoff and resolution in agreement with DFT-FDM phonon properties was 8.17 meV/Å, while others reported values as low as 25.5 meV/Å [23]. The well-tuned predictive capabilities of the SDNNFF is a result of the direct training on forces, allowing the use of 3D-distributed radial density functions in contrast to arguably more complex descriptors such as a simultaneous feed of radial and angular symmetry functions. However, the SDNNFF, like any other machine learned model for atomic interactions, requires both the generation of the *ab initio* data and the training of the neural network before evaluating the force constants, which is arguably slower than by conventional DFT-FDM for the simple periodic systems like the ones studied here. We recognized that the SDNNFF method would be more valuable in situations where FDM requires a higher number of displaced structures than that required for training and DFT becomes computationally expensive. ALM is one example that requires the evaluation of a potentially high number of randomly displaced structures for converged force constants. Another example is the calculation of higher order force constants (i.e. 4^{th} order) which exponentially increases the required number of supercells for accurately capturing the thermal conductivity of systems such as BAs, although we do not explicitly study the 4^{th} order IFCs. Here, we found the phonon properties of simple crystals to show the effectiveness of the SDNNFF in reproducing these

properties from electronic structure calculations and experiment. Once the training data is generated and the network is trained, the speed for calculating the properties is significantly faster than that of DFT by the order of 10^3.

Currently, the SDNNFF detects atomic environments purely based on spatial configurations, i.e. no chemical information is included in the inputs. Based on this, we were curious how the performance of the SDNNFF was acceptable for materials with more than one atom type, i.e. SiC and BAs, since each species may contribute to the central atom forces distinctly. We decided to analyze the SiC network in further detail by utilizing a recent tool in machine learning, t-distributed stochastic neighbor embedding (*t-sne*) [63]. While details about *t-sne* can be found in the referenced paper, in-short a *t-sne* plot reduces high-dimensional data into a 2D/3D set of points, where the proximity of the points defines their correlation (i.e. clustered points represents high similarity). Because we wanted to observe how the neural network was “thinking” in response to the local atomic configurations to produce the forces, we applied *t-sne* to the final hidden layer (i.e. before the force output) containing 500 nodes given the SiC testing set. We provide the resulting plot in Figure 2.14, which contains two significant clusters of points. We found that each cluster corresponds to configurations with central atoms corresponding to an atom inside the primitive cell, one Si and one C atom, which is intuitive since the primitive cell is the smallest unit needed to reproduce lattice symmetry and its unique atomic environments. For comparison, we also generated a *t-sne* plot for monoatomic Si

(not shown) and again found two clusters, confirming our observations about the unitcell representation trained into the SDNNFF.

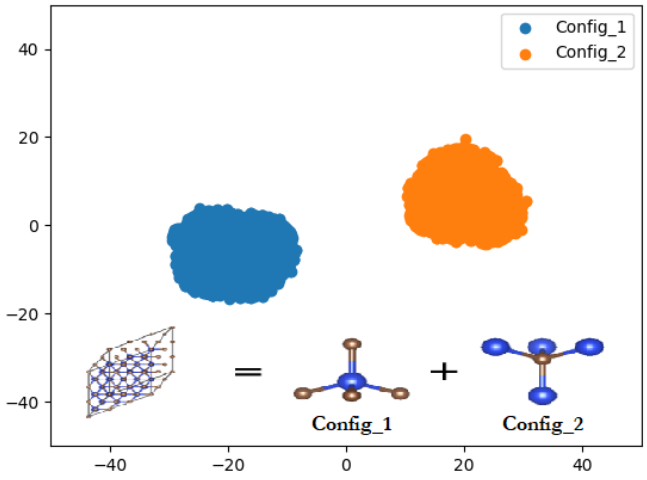

Figure 2.14 *t-sne* plot of the last hidden layer in the SiC model applied to the entire dataset.

From these results, we were led to believe that the SDNNFF was able to successfully reproduce the atomic forces for materials like SiC and BAs because the chemical environment is *consistent* with the structure, where each atomic configuration is always comprised of specific elements during training. For example, if we were to provide a set of inputs to the neural network corresponding to a central Si atom, then the network would recognize that the second neighbors correspond to C, the third neighbors with Si, etc. and correctly identify the forces associated with the values in one of the two *t-sne* clusters. We realize that each density function in space is consistently receiving the *same signals* from *equivalent atoms* in the atomic neighborhood, so the neural network learns indirectly what chemical species is contained in each density function.

Based on this observation, we expect the performance of the SDNNFF will degrade when local structure and species become independent (i.e. in alloy or multispecies amorphous material). In this case, each *t-sne* cluster would only preserve the local structure and contain a mixture of unknown elements, thereby further confusing the network. We are currently working on inserting chemical information into our density functions for future use in complex atomic environments.

The primary difference between the SDNNFF and other methods like HDNNP is the training on the atomic forces rather than the total energy. This brings about an advantage when training neural networks since the atomic forces are per-atom quantity, unlike the total energy which is a global value. Here, the DFT data for network training scales with $N \times t$, where N is the number of atoms in a supercell and t is the number of structures to be evaluated. In our case, we had 7500 structures from AIMD to be evaluated by static DFT for each material, and for a 3×3×3 primitive diamond supercell with 64 atoms, the resulting dataset contained $7500 \times 64 = 4.8 \times 10^5$ atomic configurations per material. This is unlike training a network on the total energy, where the dataset only scales with t (one total energy per structure) and therefore may require additional DFT calculations to generate a sufficiently large dataset for optimal neural network performance. We recognize other recent works that have provided schemes to take full advantage of the data from DFT, such as the extraction of and training on atomic energies from DFT by Huang *et. al.* [41], the inclusion of both total energies and atomic forces in the loss function in training the HDNNP by Minimitani *et. al.* [23],

and the additional inclusion of stress tensor by Marques *et. al.* [22]. However, in the latter two cases, atomic quantities such as the force are still treated as a single update in the HDNNP since the loss function involves an *averaging of forces* across all atoms, which is contrary to our loss function in Eq. (6) where the network is updated based on each *individual* atomic force [13]. This becomes a drawback for network training in networks like the HDNNP since vital information for the force minimization is lost with an update through one single average and the new direction for the weights/biases during training becomes less precise. In contrast, the SDNNFF preserves the $N \times t$ nature by maintaining a loss function dependent solely on each atomic force vector, and the network learns from a more robust set of examples. In addition, our networks require no additional post-processing of the DFT data since the forces are directly output from DFT, and no intermediate calculations for the loss function are necessary since Equation 2.6 only has dependence on the output forces. In other words, the training on forces in the SDNNFF not only increases the dataset size obtained from DFT calculations by explicit and independent training on each individual case, but also serves as a potential computational cost savings for training data generation and MD force calculations.

In addition to the force accuracy, another advantage of the SDNNFF is that it is an “out-of-the-box” method, where the description of the atomic environment can be controlled by only three parameters: the grid resolution, the atomic cutoff, and the local cutoff factor. As mentioned previously, while we recommend performing a convergence test for the cutoff factor D,it was observed

that a converged *D* value in one material is highly transferable to other structurally equivalent materials as seen in Figure 2.3. This makes the training of the forces very intuitive by reducing the network search to the other two parameters, the grid resolution and the atomic cutoff. Compared to the $2\times i$ parameters for the type 2 radial symmetry functions plus $3\times j$ parameters for the type 4 angular symmetry functions along with the atomic cutoff in the HDNNP method (*i* and *j* are the number of radial and angular symmetry functions respectively, also to be determined), the simplicity of the SDNNFF makes the generation of DFT force fields potentially less laborious than some existing methods [16].

Further comparing with the HDNNP, the disadvantage of the SDNNFF is that because of the equal spatial distribution of density functions in the atomic environment, a larger number of inputs may be required to cover all the important features of the atomic forces. For Si, we used at minimum 360 nodes ($k=10$) to accurately predict its thermal conductivity, while Minamitani *et. al.* obtained competitive results with only 41 inputs using the HDNNP [23]. Because the HDNNP owns more tunability, the method can obtain the feature space more efficiently than SDNNFF, requiring a smaller number of descriptors for the neural network to identify atomic configurations. Additionally, calculation of NNP inputs requires $M\times N$ distances computed, where *M* is the number of radial (symmetry/density) functions and *N* is the number of atoms inside the cutoff, which is a disadvantage for the SDNNFF since *M* scales with the number of points in *xyz* space and not radial space (i.e. calculating distances between

(±*x*,±*y*,±*z*) points and all neighbors vs. radial points with respect to the central atom and all neighbors). However, we should point out that the discussion about *M*×*N* points neglects the evaluation of additional angular symmetry functions, which is not needed in the SDNNFF since the atomic information is already represented in cartesian space. Overall, the customizability of the feature space in the HDNNP requires an increased effort to optimize its handful of parameters and is slowed down by the required total energy to evaluate the atomic forces, but on the other hand the evaluation of the HDNNP itself is quicker due to the smaller number of inputs and required neighbor distances to calculate. Therefore, at this stage it is difficult to judge which method is faster for application in MD. For now, however, we can say that the SDNNFF is capable of directly predicting the atomic forces and the resulting phonon properties of various materials with DFT accuracy, providing an emphasis on force accuracy via inclusion of 3D descriptors and inheriting a more straightforward implementation strategy for representing the atomic environment.

In comparison to classical potentials, which provides both the per-atom energies and the forces, the SDNNFF presented herein can only provide the atomic forces so far. However, we decided to use the training, validation, and testing data generated and made available by Huang et. al. to predict both forces and atomic energies of amorphous silicon using our SDNNFF method [41]. Because of the flexibility of neural networks, we were able to generate a SDNNFF+E, where we output both the force components and the atomic energies. We also generated a SDNNFF with only forces to compare with the

SDNNFF+E. Briefly covering the details, we closely followed the parameters chosen by Huang et. al., including a cutoff of 6 Å, 360 input features (i.e. *k=10*), which is on the same order of magnitude as their 550 features, two hidden layers with 500 nodes, training up to 250 cycles, and a similar loss function defined as Loss = RMSE($\{E_i\}$)+α×RMSE($\{F_i\}$) (note we arbitrarily used $\alpha = 1$, as well as *RMSE* instead of *MSE*). The resulting *RMSE* from the SDNNFF+E were 47.29 meV/Å and 15.31 meV for the forces and energies, respectively, while the *RMSE* from the SDNNFF was 42.94 meV/Å, which was slightly more accurate in exchange for no atomic energy information. In comparison to the single atom neural network potential (SANNP), the SDNNFF+E obtained an unprecedented ~50% and ~70% improvement in the forces and atomic energies, respectively. We realized that the method presented here was flexible for predicting atomic energies, although we should note that the energy is not completely invariant to spatial transformations due to the 3D representation of atomic environments. Future work should determine whether this significantly improved *RMSE* justifies the inclusion of rotationally variant energy into the SDNNFF for, e.g., bulk Green-Kubo thermal conductivity calculations. Nonetheless, the improvement in both force and energy *RMSE* is testament for the improved description of the atomic environment via 3D density functions in contrast to radial and angular symmetry functions. Moreover, we recognize that the evaluation of thermal conductivity through the nonequilibrium molecular dynamics (NEMD) such as Müller-Plathe method is possible with only the forces available, since the atomic kinetic energy exchange is used for launching the heat flux in the Fourier's law and does not

involve the atomic potential energy [64]. Overall, further study is required to bridge neural network force fields with thermal conductivity calculations in MD, with the intent to model the conductivity including all levels of phonon anharmonicity.

2.9 Conclusion

We have developed and trained a neural network force field, dubbed “SDNNFF”, capable of reproducing DFT forces and phonon properties of materials by transformation of the local atomic environment through density functions. Here, each atom was surrounded by a grid containing equidistant points within a cutoff radius. We then measured the local density of atomic neighbors at each point via a summation of radial cosine functions across all neighbors. The resulting set of density function values across all grid points led to distinguishable atomic environments for neural networks to successfully predict DFT-level force components as well as phonon dispersions and lattice thermal conductivities.

In the realm of modeling interatomic interactions through machine learning models, three achievements are accomplished in this paper. Firstly, our neural networks are directly trained on the DFT atomic forces without the need to compute derivatives from the total energy. The removal of intermediate calculations for the atomic forces simplifies the method, where a single neural network can directly predict atomic forces without the need to impose the chain rule involving the total energy and the atomic positions. Without the total energy, a direct 3D representation of the environment provides a more explicit description

of force fields in comparison to that from radial and angular descriptors, as shown from our consistently improved force *RMSE*. If the atomic energy is needed, the SDNNFF can output this along with the forces, although the calculation for the atomic energy is relatively new and, according to our knowledge, not yet standardized in DFT packages such as VASP. Additionally, the treatment of the atomic neighbors in 3D prevents the SDNNFF from preserving the rotational invariance of the atomic energy, although the improved accuracy of the SDNNFF may or may not generate negligible differences in energy. Secondly, training on atomic quantities such as forces allows the dataset to scale largely with $N \times t$, the atomic population of the structure times the number of evaluated DFT steps. Training on such tremendously large dataset strengthens the prediction capabilities of neural networks for the detection of a wide range of atomic structures one may encounter during dynamics. Consequentially, maximizing the data collected from costly DFT calculations by training the forces on a case-by-case basis becomes advantageous for neural network training. Thirdly, we found a significant reduction in human/computational effort in the search for converged neural network force fields by representation of the atomic environment with only three parameters: the atomic cutoff radius, the grid resolution, and the local density cutoff factor. The cutoff factor D may own transferability with similar structures and thus even reduce the network search down to two parameters, but transferability needs to be shown in the future for structures other than diamond structure (i.e. BCC, hexagonal, etc.).

We conclude this paper with a brief discussion on future applications of the SDNNFF. While the SDNNFF is capable of computing phonon properties with DFT-level accuracy, the speed of the SDNNFF is not utilized to its fullest potential. As mentioned before, for the above simple systems the calculation of the phonon properties with DFT-FDM is much faster than SDNNFF-ALM since the latter requires one to generate DFT data, train the network, and finally fit the force constants with the evaluated random structures before computing the phonon properties. The SDNNFF is more beneficial in applications where large systems, in particular inhomogeneous systems like amorphous materials or interfaces, and/or long timesteps are required for force evaluation, which cannot be handled by DFT due to unbearable computational cost. In other words, with a well-trained SDNNFF, large scale quasi-electronic MD simulations become possible. For phonon calculations, the SDNNFF is also beneficial for evaluating higher order force constants such as the 4^{th} order, which is necessary to capture lattice thermal conductivity of materials like BAs. The DFT-level force predictions from neural networks such as the SDNNFF are critical for the analysis of micro-scale systems where direct DFT becomes unreachable and classical MD falls short in realistically modeling these systems. In addition to thermal conductivity, applications of SDNNFF in MD simulations also include the evaluation of other dynamical properties such as diffusion coefficients and ionic conductivity, mechanical deformation, and other observables such as radial distribution function. As such, our next step for the SDNNFF is implementation into MD simulation package (e.g. LAMMPS) and investigating large-scale dynamics of

systems of interest, such as lithium ion diffusion in crystalline and amorphous silicon [65], or the bulk thermal conductivity of materials and interfacial heat transfer via Müller-Plathe method, at improved speeds by directly computing DFT-level forces, as well as finding other ways to improve the computational efficiency of the method (e.g. the reduction of network inputs or the explicit consideration of atomic species in the neural network as seen in recent studies [66,67]).

Acknowledgements

A.R. is financially supported by the NASA South Carolina Space Grant Consortium for the Minorities in STEM fellowship and the DOE fellowship for the Integrated University Program (IUP). Research reported in this publication was supported in part by the NSF (award number 1905775) and SC EPSCoR/IDeA Program under NSF OIA-1655740 via SC EPSCoR/IDeA 20-SA05. We deeply thank Dr. Emi Minamitani for her insight on the application of neural networks in atomic simulations, and Dr. Atsushi Togo for his explanations on ALM and help with compiling phonopy/3py containing the ALM package. Personal thanks from A.R. to Dr. Guangzhao Qin (Hunan University, China) for fruitful discussions on DFT calculations of the structures.

Chapter 3 Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials[2]

[2] Rodriguez, A., Lam, S., and Hu, M., 2021, "Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials," Submitted to ACS Applied Materials & Interfaces, 09/16/2021.

Abstract

Molten salts have attracted interest as a potential heat carrier and/or fuel solvent in developments of new Gen IV nuclear reactor designs, high temperature batteries, and thermal energy storage. In nuclear engineering, salts containing lithium and fluoride-based compounds are of particular interest due to their ability to lower melting points of mixtures and their compatibility with alloys. A machine learning potential combined with a molecular dynamics study is performed on two popular molten salts, namely LiF (50% Li) and FLiBe (66% LiF – 33% BeF_2), to predict thermodynamic and transport properties, such as density, diffusion coefficients, thermal conductivity, electrical conductivity, and shear viscosity. Due to the large possibilities of atomic environments, we employ training using Deep Potential Smooth Edition (DPSE) neural networks to learn from large datasets of 141,278 structures with 70 atoms for LiF and 238,610 structures with 91 atoms for FLiBe molten salts. These networks are then deployed in fast molecular dynamics to predict thermodynamic and transport properties that are only accessible at longer time-scales, and are otherwise difficult to calculate with classical potentials, *ab initio* molecular dynamics, or with experiments. The prospect of this work is to provide guidance for future works to develop general machine learning potentials for high-throughput thermophysical database generation for a wide spectrum of molten salts.

3.1 Introduction

Molten salt reactors (MSR) first appeared in the mid-twentieth century and have recently been accepted as a candidate Gen IV reactor competing with existing light-water and boiling-water reactors [68–70]. The design of MSRs centralizes molten salts as fuel and heat carriers in the primary loop, providing inherent advantages over commercial technologies such as meltdown mitigation though passive cooling, negative temperature coefficients of reactivity, low operating pressures, increased operating temperatures for efficiency, and improved waste management [71,72]. However, as molten salts are required to satisfy a diverse number of nuclear, physical, and chemical criteria to operate effectively, knowledge on salt properties is essential for MSR design [73,74]. So far, thermodynamic and thermophysical databases for molten salts are limited by the difficulty of high temperature, toxic, and corrosive experiments [75,76]. Furthermore, the possible number of salt configurations, including single, binary, ternary, and more complex mixtures of compounds, with the consideration of impurities from chromium alloys and other particulates, further increases the design space of MSRs. This huge design space calls for the development of theoretical and/or computational methods to speed up the present bottlenecks in experimental methods.

In recent years, computational simulations of candidate molten salt reactor fuels permitted a deeper understanding of their behavior and structure. Salanne *et. al.* developed the polarizable interaction potential (PIM) to study ionic liquids by inclusion of polarization effects, which has been used to study a wide variety

of single and binary molten salts involving LiF, NaF, KF, BeF_2, ThF_4 end-members with high accuracy relative to experiment [9,77–82]. Nam *et. al.* studied the structural effects of chromium dissolved in FLiBe and FLiNaK melts using *ab initio* molecular dynamics (AIMD) to observe chromium fluoride formations and diffusivity as a function of chromium valence [83]. Similarly, Lam *et. al.* studied FLiBe and FLiNaK with AIMD to observe tritium complexes and diffusivity [84]. However, the costly performance of *ab initio* methods with system size creates difficulty when predicting properties where long-time scale statistics are needed, and the *a priori* nature of classical potentials such as PIM requires the re-fitting of new potentials when encountering new end-members and/or elemental impurities in the system. Recently, a new class of potentials called the machine learned potential (MLP), allows researchers to approach materials with an 'out-of-the-box' mindset when compared to classical potentials, requiring no knowledge of the mathematical form of the potential for a particular material. Further, MLPs can reduce the severity of cost-accuracy trade-off associated with classical potentials by finding more optimal functions for describing the interatomic interactions in systems of high complexity (structural, compositional, chemical). Recently, LiF and FLiBe salts have been studied using a neural network-based MLP for the equation of state, surface energies, local structure, and self-diffusivity [85], but no works have been published that predict transport properties of molten salts with MLPs to the best of our knowledge.

In this paper, we train an MLP using the Deep Potential Molecular Dynamics (DeePMD) method to model the transport properties of single salt

lithium fluoride (LiF) and binary salt FLiBe (66% LiF – 33% BeF_2) [86]. While the previous paper by Lam utilized a modified Behler-Parinello descriptor, we chose DeePMD for its reduced manual input in hyper-parameterization. The model requires only the finite distance cutoff and max number of neighbors found in the dataset. This is a critical feature for general modeling of molten salts considering the complex and diverse local complexes formed in melts such as BeF_4^{2-} tetrahedral and polymer-like chains in denser compositions [87]. In addition, we validate the results of the training by evaluation of transport properties as functions of temperature. It is well known that the transport properties of molten salts are the most difficult to evaluate both experimentally and theoretically [6,77,88–90], and from our knowledge no other works have evaluated transport properties of molten salts with MLPs. In section 3.2, we cover the development of the LiF and FLiBe DeePMD potentials, including details on dataset generation, training, and MD. Section 3.3 details the various properties predicted from the DeepMD-Smooth Edition (DPSE) potential and compares them to other works including experimental and MD results. Finally, section 3.4 summarizes this work with a brief discussion on our future work concerning molten salts modeling.

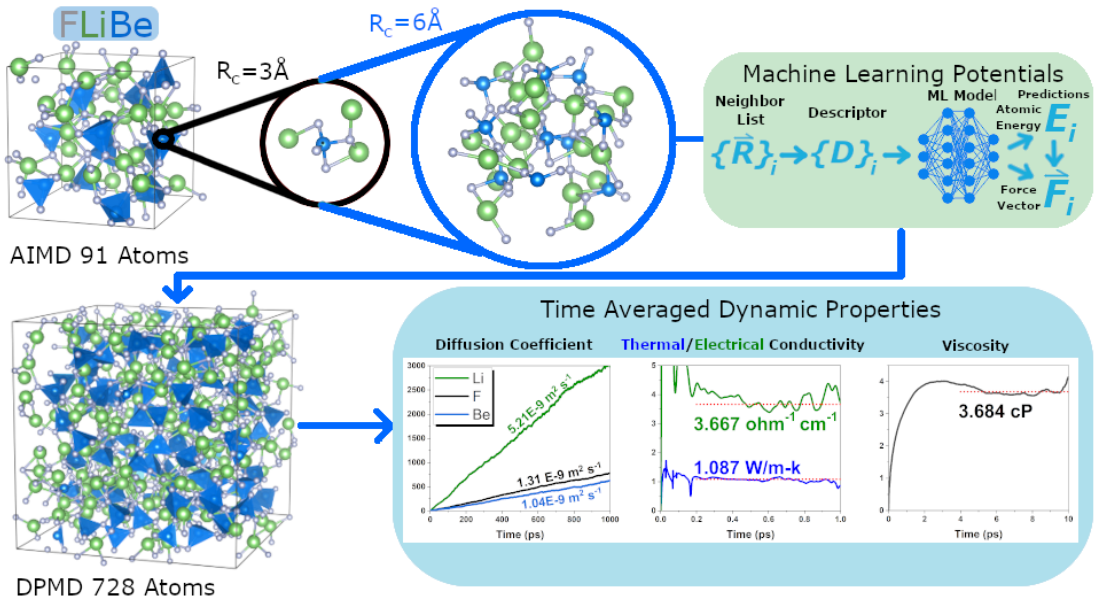

Figure 3.1 Schematic showcasing atomic simulation of molten salts.

3.2 Computational Details

Training for LiF and FLiBe neural network potentials requires a dataset that sufficiently covers the range of configurations that may manifest during finite-temperature MD simulations. As such, data for LiF and FLiBe configurations were borrowed from existing *ab initio* MD data with simulation details previously described [85]. Briefly, calculations were performed using the Vienna ab initio simulation package (VASP) with the plane-wave basis set and periodic boundary conditions in all cell directions. The plane-wave energy cutoff is 600 eV and the k-point sampling is done on a 2×2×2 Γ-centered grid. All calculations were performed using the Perdew−Burke−Ernzerhof (PBE) generalized gradient approximation (GGA) exchange−correlation functional and projector-augmented wave (PAW) potentials. The LiF dataset contains 118,115 configurations with 70 atom supercells at 1200K and experimental density, and the FLiBe dataset

contains 222,903 configurations with 91 atom supercells at 973K and experimental density. To capture changes in volume, we also generated a mixed dataset containing -5%, 1%, 5%, and 10% changes to the lattice constant for 9,208, 3,951, 5,253 and 4,751 (23,163 total) LiF configurations, respectively, and 7,168, 1,188, 4,017, and 3,334 (15,707 total) FLiBe configurations, respectively, with all systems under the same simulation parameters and conditions. With the LiF and FLiBe datasets, each were shuffled and split 80-20% for training and testing.

For training an atomic potential, we opted for the DeePMD-kit (DP-kit) package, a Python/C++ package for neural network training and evaluation of atomic energies and forces with efficient tensor operations through the Tensorflow interface [91,92]. The DeepPot-Smooth Edition (DPSE) potential contained inside the above package is chosen due to its natural symmetry conservation of atomic descriptors [93]. Here, the smooth cutoff and hard cutoff distance was chosen as 2 Å and 7 Å, respectively. The choice of hard cutoff is primarily due to time savings in FLiBe MD (1.3187×10^{-3} seconds/step/atom for 7 Å vs. 2.3187×10^{-3} seconds/step/atom for 9 Å) with negligible 0.03 g/cm^3 improvement in the predicted density, whereas the smooth cutoff from 8.8 Å to 2 Å shows slight ~0.01 g/cm^3 improvement with no notable difference in speed. The resulting training of DPSE networks yielded a testing root mean squared error of 2.54×10^{-5} eV/atom and 13.8 meV/Å for LiF energy and forces, respectively, and 4.86×10^{-5} eV/atom and 23.2 meV/Å for FLiBe energy and forces, respectively. As exhibited by the comparison of the testing energies and

forces in Figure 3.2, the excellent accuracy on untrained data suggests well-represented potentials for simulating *ab initio* potential energy surfaces and ensuing dynamics for LiF and FLiBe molten salts.

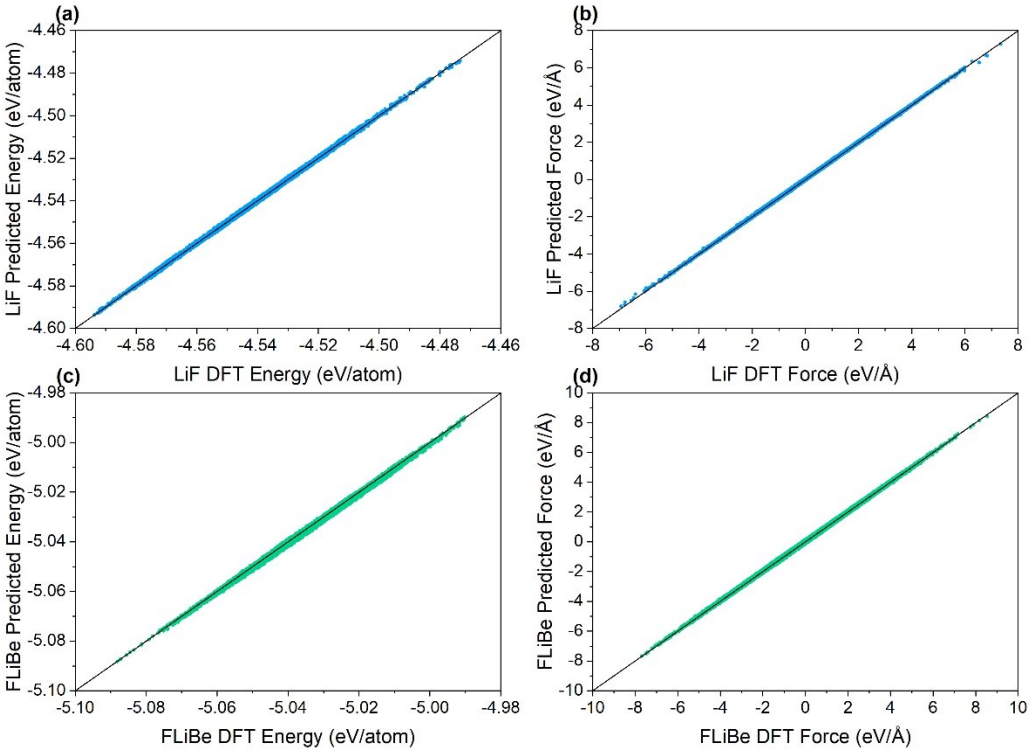

Figure 3.2 Plots comparing unobserved testing DFT and DeePMD predictions for the LiF and FLiBe energy and forces.

To generate acceptable MLPs with high quality of energy and force, sufficient data must be provided during training to capture, at minimum, a range of all possible configurations that may be found during finite temperature MD simulations. Due to the costly nature of *ab initio* calculations, it is of interest to minimize the computational cost from the data generation step by requiring the least number of configurations needed for converged force and energy testing

errors, such that any further training is not necessary to sufficiently generalize unseen structures. We have provided Figure 3.3 which contains both energy and force error against number of AIMD steps. From our existing FLiBe data, 50,000 constant volume AIMD steps were isolated as training data, leaving the remaining 172,903 structures for testing. We further split the training data into segments and train various models with increasing number of steps, each inheriting the same data as the last. After training, we evaluate the same testing set with the models and compare the model output with *ab initio* energies and forces. As seen from the figure, acceptable RMSE values are reached around 10,000 steps, with slight fluctuations in the force and energy RMSE thereafter.

Furthermore, the effect of compressed/expanded volume data on model accuracy is investigated. Provided the converged model at 10,000 constant volume configurations, we re-train new models with the same previous 10,000 steps but with the existing variable volume data, such that 1000, 2000, 4000, and 7000 compressed, expanded, and 50/50 compressed/expanded volume data are included for training (i.e. yielding 12 new networks). We then evaluate the same testing set as before to find the RMSE. Note, the expanded volume data is randomly selected from the pool of +1%, +5%, and +10% lattice constant datasets. From our training of additional volumes, all force RMSE values lie on the converged 23.5 meV/Å found from constant volume networks, while the energy RMSE fluctuates about 3×10^{-5} eV/atom, more than double of the 1.21×10^{-5} eV/atom in the constant volume case. Although an energy RMSE increase is observed with added volume data, we emphasize this as a necessary trade-off

to improve the quality of molten salt MD simulations, due to the improvement of the predicted density and subsequent properties as discussed in the next section. The networks trained with constant volume either resulted in very low or unstable densities in which excessive cell expansion occurs, with no particular pattern in the increasing number of constant volume configurations. Here, we briefly show how the inclusion of variable volumes improves the resulting density via MD simulations with the aforementioned DPSE models. The inset of Figure 3.3 shows a 10.4% improvement to the predicted density at 973K with additional volume data. Interestingly, all networks including expanded data seem to converge with only 1,000 additional structures (500 expanded structures in the compressed + expanded case), whereas the pure compression case requires at least 4,000 to reach similar density. We assign the performance increase by the inclusion of forces in an expanded system, with the DPSE model more accurately capturing the calculated pressure under cell expansion and converging to a relatively higher density. On the other hand, without the expanded data, the pure compressed case requires more samples to reach density convergence. We theorize that the compressed system improves the quality of short-ranged interactions and promotes the attractive potential between atoms in the system.

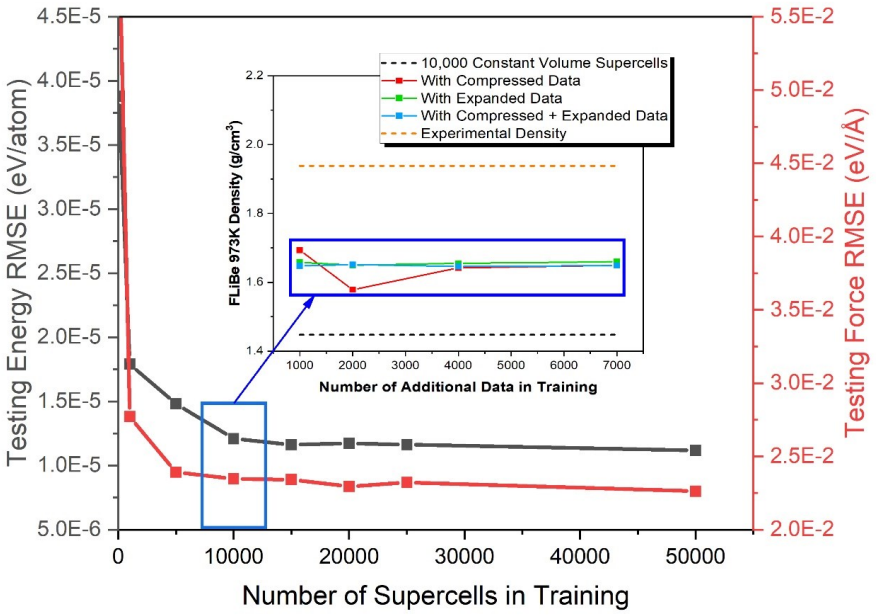

Figure 3.3 The energy and force RMSE for FLiBe against number of supercells used in training.

After training on the dataset, the DPSE potentials were used in the Large Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) via interface with DeePMD-kit [10]. Specifically, a pair style in LAMMPS uses the trained model for energy and force calculation during MD steps. Exploiting the linear scaling of DPSE, the supercell from AIMD is extended to a 2×2×2 system as the initial structure, resulting in 560 LiF atoms and 728 FLiBe atoms. The temperature was chosen such that the system is well beyond its liquidus, thus the systems are simulated at 1120K-1320K at 40K intervals for LiF and 800-1100K at 100K intervals + 973K for FLiBe. With 1 fs as the timestep, the system is first relaxed for 250 ps under NVT, followed by 250 ps under NPT, where lattice constant is taken from the average volume during NPT. Then, we perform two

sets of production runs at various temperatures, one set using the calculated relaxed volumes and the other set at experimentally known densities. Both of these sets of simulations undergo 100 ps of NVT, followed by 1000 ps of NVE. We note that the temperature range of the MD simulations is outside the constant temperature simulations of the training dataset. Nonetheless, the structure of the melts does not vary strongly within the small temperature range as seen in the next section and the current model was sufficient for our study.

To display the linear scaling of MLPs, we generated Figure 3.4 containing the CPU time per timestep against the system size of LiF. As seen from the figure, the evaluation time for AIMD scales closely with the number of atoms cubed. On the other hand, the rigid ion model (RIM) [94,95], a parameterized potential, the Spatial Density Neural Network Force Field (SDNNFF) [96], a machine learning force field, and DPSE all scale linearly with the system size. Although the ML-based methods are more time consuming in comparison to parameterized potentials, they offer the same linear scaling quality inaccessible by first-principles while representing the potential energy surfaces seen in DFT. Moreover, the computational cost of MLPs can be reduced by a constant if efforts are made to prune the model or systemically reduce the number of inputs to remove redundant or insignificant parameters, after the network has been trained. In this way, the MLP performance can be tuned based on the required sensitivity/accuracy of the system in a way that classical models cannot.

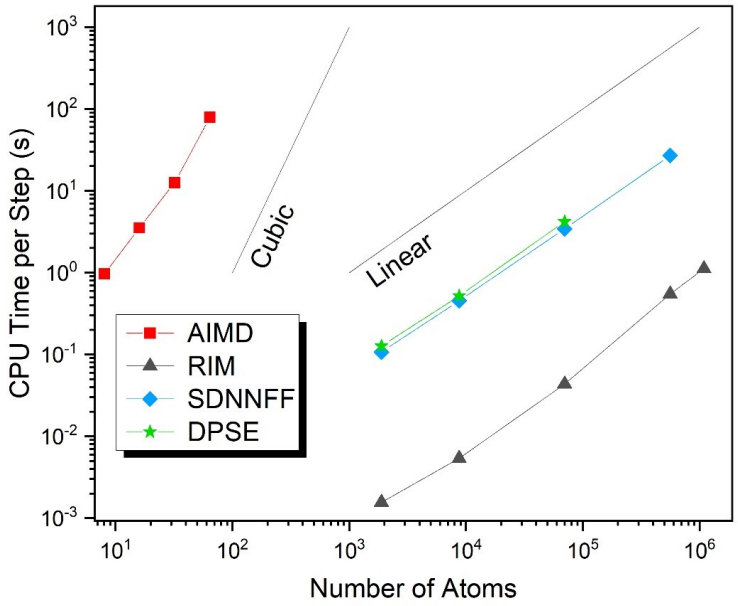

Figure 3.4 Time scaling of various atomic simulation paradigms against system size.

3.3 Thermodynamic and Transport Properties

For determination of transport properties in MD, the density ρ of molten salts is closely related to composition and existing local structures in the salt and is essential for matching subsequent properties with experiment. Several studies for the thermal conductivity through equilibrium MD simulations have shown direct correlation with the number density of the melt and cell volume [82,97]. Additionally, the model developed by Gheribi *et. al.* for thermal conductivity prediction strongly depends on known densities of molten alkali and alkaline earth salts, with a relatively weak dependence on the temperature [88]. As mentioned in the previous section, NPT relaxations on LiF and FLiBe were performed to obtain the equilibrium volumes from the DPSE potential and the

resulting densities versus the temperature are plotted in Figure 3.5. For Figure 3.5.a, in order as shown in the legend, the density for LiF under NPT are obtained from our work, MD with the PIM model [6], and experiment [98]. Similarly, in Figure 3.5.b, values for FLiBe under NPT from our work, MD with the PIM model [99] and experiment are shown [98]. As seen in the figure, the densities follow a similar downward trend as experiment but is ~14% lower when comparing the absolute value. We attribute the higher average volume to the quality of the DFT dataset the DPSE was trained on. A previous study by Nam *et. al.* determined by first principles MD that the effect of van der Waals (vdW) on FLiBe volume is significant, such that no consideration of dispersion forces can lead to 18% higher volume with the PBE pseudopotential [83]. Although the addition of mixed volumes in the training set promotes higher density as seen in the inset of Figure 3.5, the resulting densities are still within the lower 10%-20% region which we attribute to the lack of dispersion interactions in the training data. Moreover, accurate inclusion of dispersion effects can be non-trivial. In standard deep neural network potentials and DeepMD, distance cutoff results in truncation of long-range electrostatic and dispersion interactions. Recent efforts (third and fourth generational NNIPs) have sought to resolve this using additional neural networks to predict environment-dependent charges that can be used to predict long-range interactions [100]. For this work, we proceed to evaluate other transport properties in MD at both NPT relaxed volumes and experimental densities to emphasize the importance of capturing density, a relatively straightforward property to calculate, before computing complex transport

properties, such as the thermal conductivity. For future studies, we strongly suggest including vdW dispersion interactions such as vdW-DF or DFT-D2 in generating molten salt datasets from DFT, since exchange-correlation functionals cannot consider these dispersion interactions on their own. Further, we recommend testing different neural network schemes for representing these interactions with high fidelity.

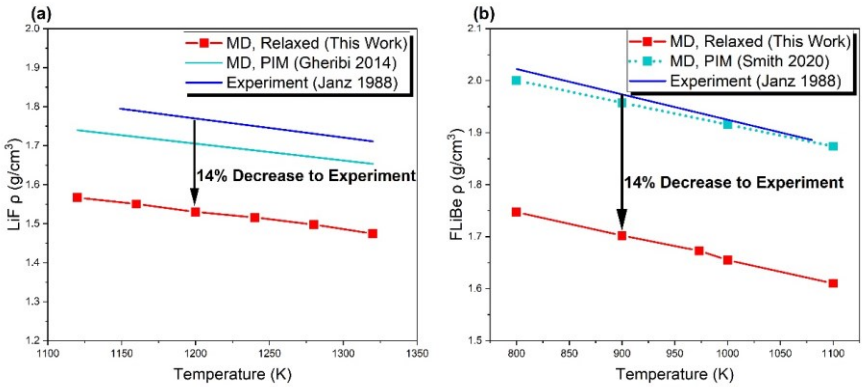

Figure 3.5 Plots of the density against temperature for LiF and FLiBe above the melting point.

To measure the local structure of ionic fluids, the radial distribution function (RDF) and diffusion coefficients are compared between DPSE-MD, AIMD, and results from other papers. As seen from Figure 3.6, the RDF curves found from DPSE-MD match excellently with AIMD and confirms the efficacy of the potential. Furthermore, these curves for FLiBe and LiF offer insight to the decrease in density after MD relaxation. For the inset of Figure 3.6, only the Be-Be RDF for the relaxed structure is compared with that from the experimental structure for brevity since the other RDF pairs show little variation in peak

positions. As such, the drop in density is attributed to the expansion between Be-Be pairs, as noted by the stretched inter-peak distances for relaxed structures. This phenomenon has also been observed by Jabes *et. al.* when comparing the transferable rigid ion model (TRIM) and polarizable ion model (PIM) atomic potentials for FLiBe, where lack of polarizability effects in the TRIM potential led to stretched Be-Be peaks and a reduction in density by 17-20% [87]. The expansion of RDF peaks is also observed for the LiF system as seen in the inset for Figure 3.6. Again, we isolated the F-F pair for brevity which contains the largest discrepancy between relaxed and experimental volumes. Similar behavior is found as with the Be-Be pair for FLiBe, with relaxed LiF experiencing larger F-F peak distances although the stretch is not as pronounced as that for the FLiBe Be-Be pair. Due to the shared volume expansion of both systems, the abscense of dispersive interactions from AIMD results in a weak representation of the interactions between like-charged ions, which are essential for the precise modeling of local structures and the resulting density in molten salts.

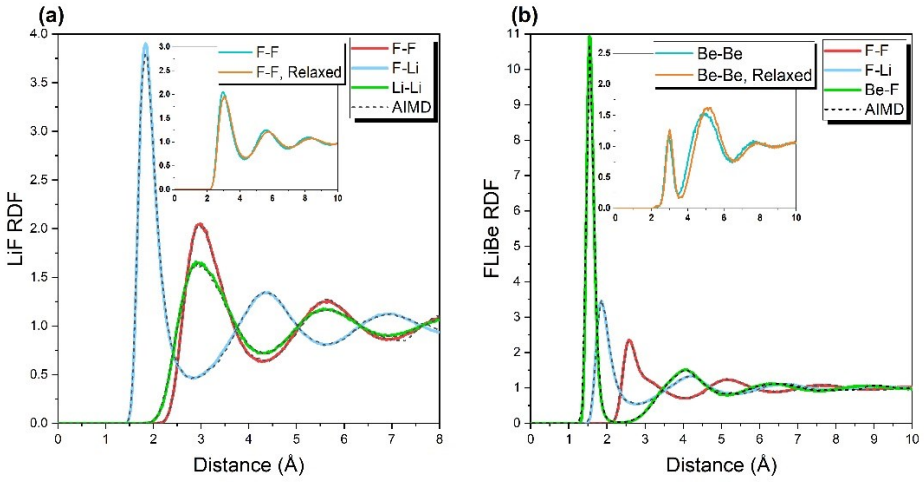

Figure 3.6 RDF curves for LiF at 1200K and FLiBe at 973K in comparison to AIMD.

To showcase the temperature dependence on the local structures of liquid LiF and FLiBe, Figure 3.7 plots RDF curves for various temperatures found from DPMD. The largest observable discrepancy is in the height of the first peak positions of all curves for which the highest peak corresponds to the lowest temperature simulation here (1120K for LiF, 800K for FLiBe), decreasing with temperature increase. Although at low temperature the peak size is maximized, the radial position of the peak itself experiences no changes across all temperatures, which suggests an approach to crystallization. Similar behavior has been shown for LiF with the Born-Mayer potential, wherein first peak positions near the captured melting point closely correspond with those at much higher temperatures, meanwhile solid LiF differs in both peak positions and height from its liquid counterpart [101]. Provided, LiF and FLiBe molten salt data at constant temperature for training captures the local atomic structures at

surrounding temperatures, since the structures in liquid phase are not sensitive to changes across their respective temperature ranges. Additionally, the aforementioned weak dependence of molten salt densities on the temperature (i.e. -0.1 g/cm^3 from 800-1100K in FLiBe experiments) also supports the relatively small changes in local structures of the melt. However, we warn others that this behavior may not transfer to other salt mixtures for all temperatures in the liquid phase, and testing for local structures is recommended to discover those unaccounted for at other temperatures especially near the melting point.

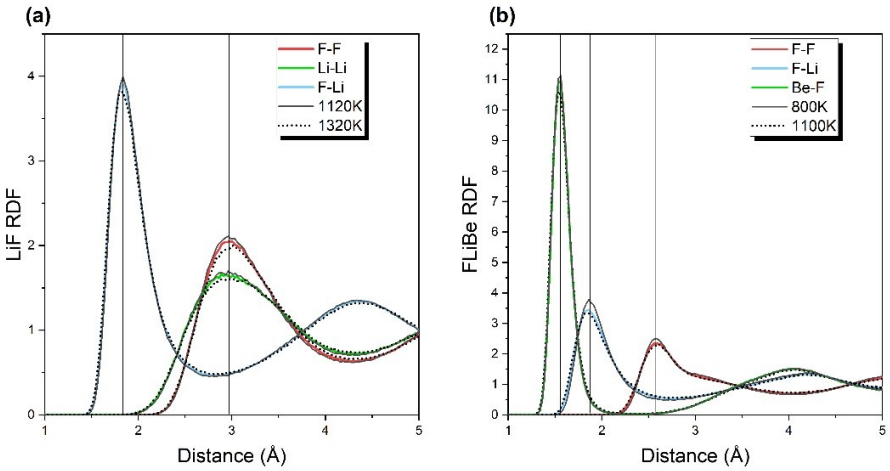

Figure 3.7 RDF curves against temperature for LiF at 1200K and FLiBe at 973K.

In tandem with the structure, the dynamic behavior of molten salts also follows closely with the density and is measured by the diffusion coefficient *D*. Here, the diffusion coefficients of ions are calculated through the slope of the mean square displacement *via* the Einstein relation,

$$D = \frac{1}{6}\lim_{t\to\infty}\frac{\mathrm{d}}{\mathrm{dt}}\left[\frac{1}{N\times n_t}\sum_{n_t}^{j=1}\sum_{N}^{i=1}\left(r_i(t_j+dt)-r_i(t_j)\right)^2\right] \quad (3.1)$$

where the term inside the brackets is the mean square displacement (MSD), averaged over atoms N of type i and subsets n_t of time t_j [85]. In Figure 3.8.a, diffusion coefficients for LiF at experimental and NPT relaxed volumes are obtained from our work, MD with the PIM model [79], and experimental results [79]. Similarly, for Figure 3.8.b, diffusion coefficients for FLiBe at experimental and NPT relaxed volume are obtained from our work, MD with the HDNNP, and AIMD results [85]**.** The diffusion coefficients in Figure 3.8 increase with the temperature, with lithium ion diffusion dominating both LiF and FLiBe systems. Additionally, the fluorine ion diffusion in LiF is significantly more diffusive than that in the FLiBe case, where fluorine diffusion is much closer to the beryllium ion diffusion. The formation of tetrahedral BeF_4^{2-}, $Be_2F_7^{3-}$, and more complex polymer chains may be attributed to the low fluorine and beryllium ion movement due to the increased effective mass of these complexes [84]. Meanwhile, the lithium ion in FLiBe experiences no polymerization and maintains its high mobility. For comparison, the plots also display the diffusion coefficients for the relaxed systems, consistently yielding larger values albeit the MD temperatures were equivalent to those at experimental volumes. In fact, the LiF and FLiBe diffusion coefficients for the lower density case follow extremely similar trends as those at experimental density, implying no qualitative changes to the structure. This behavior is indicative of a more free-flowing system, where ions and Be-F complexes are less likely to scatter and maintain higher average velocities in the larger simulation cell. Finally, excellent agreement of the DPSE-MD results with

other works are shown, although not many points are plotted since current literature is starved of diffusion data due to the difficulty in measurement [75].

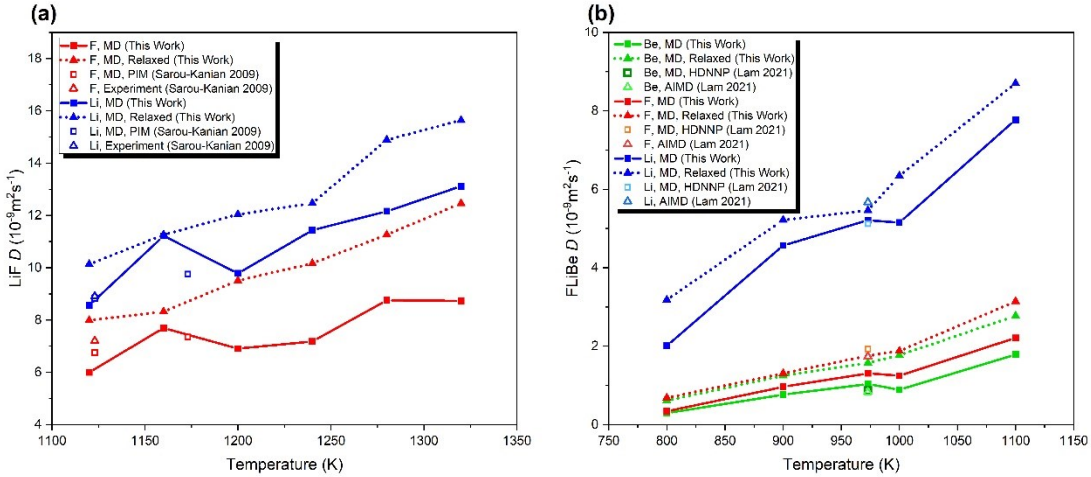

Figure 3.8 Plots of the diffusion coefficients against temperature for LiF and FLiBe above the melting point.

The thermal conductivity λ for ionic liquids is computed in equilibrium molecular dynamics by Green-Kubo relation through integration of autocorrelation between heat and charge fluxes through the phenomenological coefficients $L_{\alpha\beta}$ [102,103],

$$L_{\alpha\beta} = \frac{1}{3k_B V}\int_0^{\infty} <\boldsymbol{j}_{\alpha}(t)\cdot\boldsymbol{j}_{\beta}(0)>dt \tag{3.2}$$

where thermal conductivity is defined as [99]

$$\lambda = \frac{1}{T^2}\left(L_{ee} - \frac{A}{B}\right) \tag{3.3}$$

For single component molten salts, $A = L_{ez}^2$ and $B = L_{zz}$ where e and z refer to the heat and charge terms, respectively, and for binary salts,

$A = L_{ez_1}^2 L_{z_2z_2} + L_{ez_2}^2 L_{z_1z_1} - 2L_{ez_1} L_{ez_2} L_{z_1z_2}$ and $B = L_{z_1z_1} L_{z_2z_2} - L_{z_1z_2}^2$. In our simulations, ion 1 is the beryllium ion and ion 2 is the lithium ion. The formulism for charge flux in the binary case used in this work is [82]

$$\boldsymbol{j}_{z_n} = \sum_n (q_n - q_f) * \boldsymbol{v}_n \tag{3.4}$$

with n being over all cations in the melt and f as the anion ($q_f = -1$). For improved convergence of the autocorrelation in Equation 3.2, we divide the 1-ns simulation into 1000 1-ps samples and perform a time averaging of the resulting thermal conductivities over these samples.

Because the thermal and electrical conductivity of ionic liquids requires the charge flux of atoms in the simulation cell, we assigned fixed charges corresponding to their chemical element, i.e. Li with +1*e*, F with -1*e*, and Be with +2*e* charges. However the LAMMPS simulations do not consider additional coulombic forces. This is because the DPSE, like other second generation ML potentials, invoke full electronic interactions including those of coulombic nature up to the cutoff, but no long-ranged interactions or charge transfer between atoms are considered [100]. Regardless, we obtain agreeable results for the thermal and electronic conductivity in the systems studied here, possibly due to the strongly ionic, well-defined character of the ions in the melt. Future works on MD simulations of molten salts including fixed charges, dynamically altered partial charges, such as those found in the modeling of HfO_2 RRAM cells using the ReaxFF potential [104], or MLP-trained with partial charges from charge

equilibriation method [100] combined with long-ranged solvers for coulombic interactions should be considered.

The thermal conductivity found from DPSE-MD and other works are depicted in Figure 3.9. For Figure 3.9.a, in order as shown in the legend, thermal conductivities for LiF at experimental and NPT relaxed volumes are obtained from our work, MD with the polarized ion model [99], MD with the rigid ion model [97], fitted line from MD with polarized ion model [6], experimental results [105,106], and by a predictive model [88]. Similarly, in Figure 3.9.b, thermal conductivities for FLiBe at experimental and NPT relaxed volume are obtained from our work, MD with the polarized ion model [99], experimental results [107,108], and by a predictive model [89]. Error bounds are shown in light green for MD simulations at experimental density and black for those at relaxed densities. The heights of the error bounds represent the RMSE about the average thermal conductivity at the tail-end of the autocorrelation curves. As seen for the LiF thermal conductivity, results from the simulations at experimental density fall between to those from other MD simulations and experimental results. The thermal conductivity for the relaxed case shows even further improvement although this may be coincidence from the increased ion diffusivity in larger volumes. For FLiBe, the agreement with relatively constant thermal conductivity from experiment also matches well with the DPSE results. For the relaxed case, the thermal conductivity remains below that at experimental density until 1100K, where the thermal conductivity suddenly jumps. Unlike the thermal conductivity for single component salts, the ratio of determinants *A/B* for binary

salts contains both additive and subtractive terms to the thermal conductivity, and the cause for the behavior becomes non-trivial. We further discuss the difference for thermal conductivity contributions in low and high density simulations below.

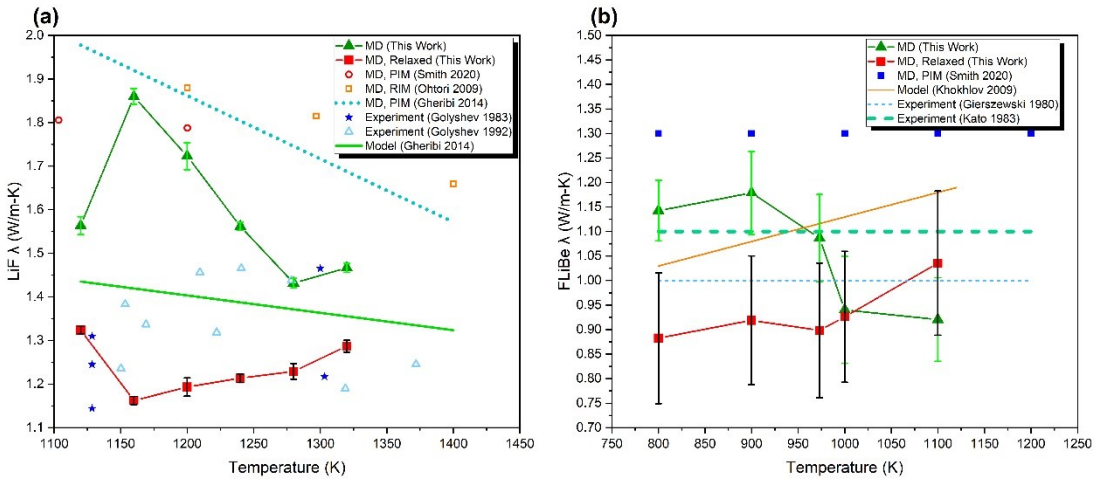

Figure 3.9 Plots of the thermal conductivity against temperature for LiF and FLiBe above the melting point.

To understand the effects of the phenomenological coefficients in Equation 3.2 on the resulting thermal conductivities of molten salts, their contributions under experimental and MD-relaxed densities are plotted in Figure 3.10. The construction of the thermal conductivity for single salts is subtractive from the pure heat flux term L_{ee} by the cross-correlation charge and heat flux ratio L_{ez}^2/L_{zz}. For LiF in Figure 3.10.a. and 3.10.b., the thermal conductivity is approximately halved from that of the heat flux term, yielding a much lower thermal conductivity than the pure heat flux thermal conductivity. On the other hand, for binary salts, the thermal conductivity contributions are no longer purely subtractive, but follow a more complex set of contributions by the ratio of determinants *A/B*. As seen

from Figure 3.10.c. and 3.10.d., while a large portion of the thermal conductivity of FLiBe originates from the heat flux term L_{ee}, an equally large portion is canceled out by the $L_{ez_2}^2 L_{z_1z_1}/B$ term. In our results, we noticed much higher values amongst the cross-correlation constants for the lithium ion charge flux with the heat flux L_{ez_2} and to itself $L_{z_2z_2}$, explaining the large conductivity contributions of subtractive $L_{ez_2}^2 L_{z_1z_1}/B$ and additive $2L_{ez_1}L_{ez_2}L_{z_1z_2}/B$. The high contribution to thermal conductivity by lithium may be assigned to its high mobility in FLiBe in comparison to the mobility of fluorine and beryllium ions. In LiF, the high mobility of both lithium and fluorine ions dampens effect on the subtractive L_{ez}^2/L_{zz} term, unlike in FLiBe where the heat flux term is essentially canceled out by the existing disparity in ion diffusivity. The lowest contributor to the conductivity, $L_{ez_1}^2 L_{z_2z_2}/B$, is primarily dictated by the square of the heat-charge flux correlation of the beryllium ion $L_{ez_1}^2$ and was observed to be about a tenth of the $L_{z_2z_2}$ term. This lowers the effective thermal conductivity reduction of the $L_{ez_1}^2 L_{z_2z_2}/B$ term and has the least effect on the conductivity. Hence, the term that maintains the conductivity of FLiBe in addition to L_{ee} is the $2L_{ez_1}L_{ez_2}L_{z_1z_2}/B$ term in which all the heat-charge flux terms are weighted similarly (i.e. power to the unity).

To analyze the effects of high (experimental) and low (MD relaxed) density on the resulting thermal conductivity with respect to temperature, their results for LiF and FLiBe are placed together for comparison in Figure 3.10. Here, solid lines are additive to the thermal conductivity, while dashed lines are subtractive.

The total thermal conductivity is also shown. In experimental density LiF, the heat flux term L_{ee}/T^2 in Figure 3.10.a. greatly influences the thermal conductivity, both experiencing an overall decreasing trend with relatively fixed $L_{ez}^2/L_{zz}/T^2$. Indeed, experimental density FLiBe also experiences decreasing thermal conductivity with temperature but is described by a more complex set of coefficients involving several cross-correlation terms with the heat flux, i.e. L_{ez_1} and L_{ez_2}, and the behavior is increasingly non-trivial when compared to LiF. In Figure 3.10.c., the decrease is primarily dominated by the increase in the subtractive $L_{ez_2}^2 L_{z_1z_1}/B/T^2$ term, most likely attributed by the sudden jump in lithium ion diffusion after 973K. Nonetheless, this inverse relation of the thermal conductivity with temperature for both single and binary salts is mentioned in previous works and has been attributed to the decreasing density of the melt [82,97]. On the other hand, the thermal conductivities of the melts for the MD relaxed volumes in Figure 3.10.b. and 3.10.d. are directly proportional to the temperature and deviate from experimental results despite owning a decreasing density trend as observed in Figure 3.5. For low-density LiF, both the L_{ee}/T^2 and $L_{ez}^2/L_{zz}/T^2$ term remain relatively constant with their difference yielding a slight increase in the conductivity. Although the L_{ee} term increases in both the low and high-density simulations, low-density L_{ee} rises more rapidly with temperature than the high-density L_{ee} as seen by the ratio with T^2 and results in a mostly temperature-independent L_{ee}/T^2 contribution. For low-density FLiBe, the increase in thermal conductivity is made possible by a decline in both subtractive $L_{ez_2}^2 L_{z_1z_1}/B/T^2$ and

$L_{ez_1}^2 L_{z_2z_2}/B/T^2$ terms after 900K, more so than the decline in the additive $2L_{ez_1}L_{ez_2}L_{z_1z_2}/B/T^2$ term. The drop in all three terms with temperature of which involve charge-charge and charge-heat cross-correlations is again most likely due to the weaker scaling with T^2 for the low-density case. We note that despite division by $B = L_{z_1z_1}L_{z_2z_2} - L_{z_1z_2}^2$, both high-density and low-density simulations owned matching values of B and contributed little to the mentioned trends. In summary, we have shown how the simulated density effects thermal conductivity behavior against the temperature, reinforcing our earlier statement to accurately capture the density before approaching other properties.

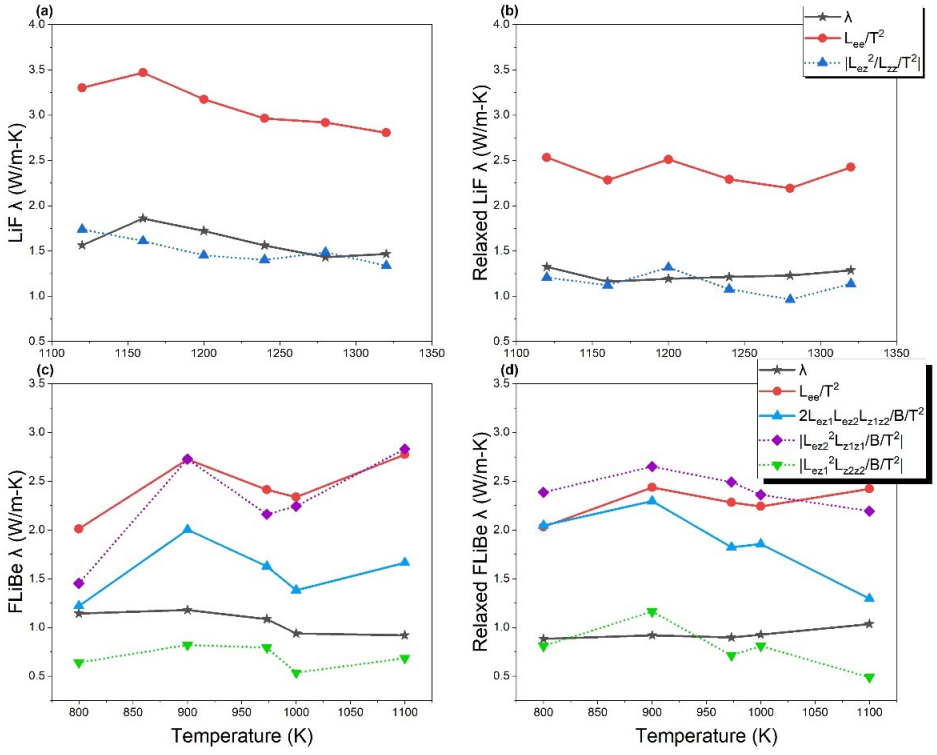

Figure 3.10 Contributions to the thermal conductivity for LiF, relaxed LiF, FLiBe, and relaxed FLiBe.

The electrical conductivity σ in MD simulations may be found through the MSD of charges in the melt, the Nernst-Einstein approximation, or through the Green-Kubo formalism [78,102]. Since the Green-Kubo thermal conductivity for ionic liquids requires the charge correlation term L_{zz}, the electrical conductivity is readily available by dividing by the temperature. Note that the L_{zz} term is found through the charge flux of all atoms in the system, hence the charge flux for single component salts is applied to the electrical conductivity for FLiBe. For improved convergence of the autocorrelation, we divide the 1 ns into 1000 1-ps samples and perform a time averaging of the resulting electrical conductivity. The results for LiF and FLiBe electrical conductivity are provided in Figure 3.11 with fitted experimental results from Janz *et. al* [98]. Here, electrical conductivities for LiF and for FLiBe at experimental and NPT relaxed volumes are obtained from our work. Both plots show experimentally fitted data, where dotted lines represent extrapolated data [98]. The insets show extensive electrical conductivity with units $10^{20}\ \mathrm{ohm^{-1}cm^{2}}$. Error bounds are shown in light green for MD simulations at experimental density and black for those at relaxed densities. The heights of the error bounds represent the RMSE about the average electrical conductivity at the tail-end of the autocorrelation curves. Here, the electrical conductivity matches well with experiment following an increasing trend against temperature. Interestingly, the results from low density simulations yield lower electrical conductivity values despite owning higher ionic mobility. The difference in the electrical conductivity is primarily attributed to the division by the volume in the L_{zz} term from Equation 3.2. The insets of Figure 3.11 contain the electrical

conductivity multiplied by the cell volume for the extensive electrical conductivity, where improved agreement between the relaxed and experimental density simulations are observed. For LiF, slight disagreement remains although based on the fluctuations of the electrical conductivity we attribute the uncertainty toward the Green-Kubo MD. For FLiBe, an improved overlap in the extensive electrical conductivity is observed. Based on these results, the electrical conductivity owns a strong dependence on the density.

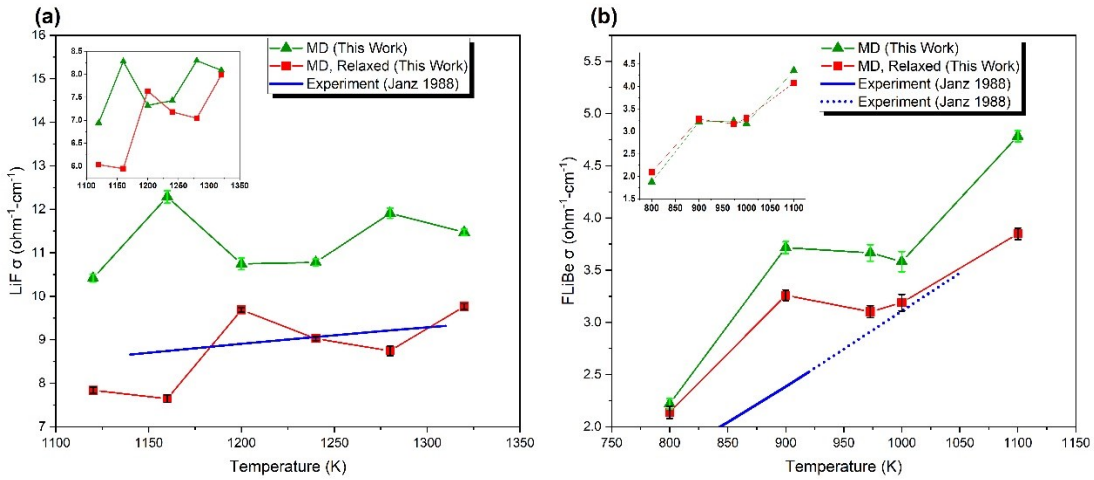

Figure 3.11 Plots of the electrical conductivity against temperature for LiF and FLiBe above the melting point.

The shear viscosity η is found similarly to the Green-Kubo thermal conductivity [109],

$$\eta = \frac{V}{k_B T}\int_0^{\infty} \langle P_{\alpha\beta}(t)\cdot P_{\alpha\beta}(0)\rangle dt \qquad (3.5)$$

where $P_{\alpha\beta}$ represents the off-diagonal terms in the pressure tensor during MD. For improved statistics, the viscosity is averaged over P_{xy}, P_{xz}, and P_{yz} terms. Optionally, the viscosity may also be averaged over $P_{xx\text{-}yy}$ and $P_{2zz\text{-}xx\text{-}yy}$ but the first three terms were sufficient for convergence of the autocorrelation function [6]. In addition, due to the arrhenius relation of the viscosity to temperature, the convergence of the autocorrelation in Equation 3.5 requires variable autocorrelation times for the time averaging over the span of the 1 ns simulation time. Specifically, for convergence of viscosity plateaus regardless of the simulated experimental or relaxed densities, the autocorrelation times for LiF were 5 ps for all simulated temperatures, while those of FLiBe were 15, 15, 10, 5, and 5 ps in order of increasing temperatures. As seen in Figure 3.12, the viscosity of LiF and FLiBe at experimental densities matches well with experimental works, while those found from relaxed supercells own significantly lower viscosity. In Figure 3.12.a, tn order as shown in the legend, viscosities for LiF at experimental and NPT relaxed volumes are obtained from our work, MD with the polarized ion model [6], and experiments [110,111]. Similarly, in Figure 3.12.b, viscosities for FLiBe at experimental and NPT relaxed volumes are obtained from our work, MD with the polarized ion model [99], and experiments [107,112]. Error bounds are shown in light green for MD simulations at experimental density and black for those at relaxed densities. Those with limited height are omitted for visual purposes. The heights of the error bounds represent the RMSE about the average viscosity at the tail-end of the autocorrelation curves. In comparison to the PIM model, the LiF viscosity agrees well with other

works, while the FLiBe viscosity underestimates experimental viscosities with PIM overestimating by a similar magnitude. It is well known that the viscosity is a difficult property to predict from MD simulations due to the required long simulation times for converged autocorrelation in Equation 3.5, and here the relatively short 1 ns simulation time is a potential source of error [75,99]. Nonetheless, the agreement with experimental trends are acceptable with matching errors from those of other MD works. Other works have shown that the viscosity owns an inverse relationship with the diffusion coefficient [101], and the diffusion coefficient increases with decreased density as previously observed in Figure 3.8. For example, the higher ionic mobility in the relaxed supercell is consistent with the lowered viscosity by further increasing the ionic mobility and decreasing the viscosity from expansion due to the temperature. As such, the viscosity is directly proportional to the density of molten salts as observed by its decrease with increasing volume and temperature.

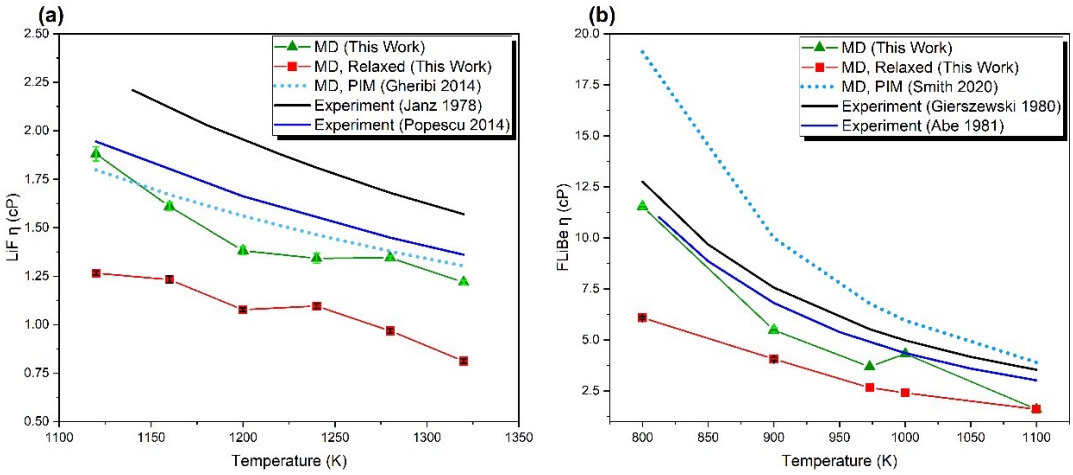

Figure 3.12 Plots of the viscosity against temperature for LiF and FLiBe above the melting point.

3.4 Conclusion

An investigation on the thermodynamic and transport properties of molten salts LiF and FLiBe is performed, where flexible and rapid machine learning potentials are fitted with DFT data without the need of pre-existing physical models. The matching of local structures with AIMD and diffusion coefficients with experiment displays the excellent ability of the machine learning potentials to reproduce complex energy surfaces and realistic dynamics. With the linear scaling of Deep Potential Smooth Edition (DPSE) neural networks, we explored the dynamics of molten salts using large supercells through MD with DFT accuracy. Due to the difficulty of evaluating dynamic properties in experiments, the diffusion coefficients, thermal conductivity, electrical conductivity, and viscosity of LiF and FLiBe were predicted within acceptable accuracy of experiments and other works using MD. Although the density of molten salts only

has reasonable agreement with experimental results, the resulting densities are linked to the lack of dispersion relations in the training data and match previous works excluding these interactions [83,87]. In future studies, dispersion considerations in DFT packages will be studied to consider polarization in the training data and improve atomic interactions in molten salt machine learning potentials. Inclusion of long-ranged interactions beyond the cutoff of DPMD may also improve the resulting density.

Here we conclude this work with a perspective on the future of atomic molten salt modeling. With the massive design space of molten salts, developing parameterized potentials e.g. PIM potentials for all possible compositions would take much effort and time. Instead, the “hands-free” approach of machine learning potentials is crux for the development of general potentials that can simultaneously cover a wide range of molten salt compositions and temperatures, essentially spanning across the periodic table. The multi-dimensional, non-linear nature of MLPs allows for a higher degree of transferability that is cut short by the limited flexibility of parametric potentials. However, one of the main limitations of recent machine learning potentials is the exponential scaling of model parameters with number of elements. For example, the popular high-dimensional neural network potential (HDNNP) requires N radial and $N(N+1)$ angular symmetry functions for each element-specific network, which can quickly slow down both training and the evaluation of structures with increasing elements N [66]. Although the DPSE potential used in this paper necessitates no user-specified descriptors such as symmetry functions, the

model still requires N fitting networks with N^2 embedding networks and is increasingly more time-consuming during training and evaluation steps for large number of elements [113,114]. Furthermore, descriptors should be designed in such a way that any number of elements may be added without needing to reconstruct and train the model from scratch. Generating such descriptors from data spanning many elements could enable a reduction in, e.g., network complexity if a new element only changes the NN input numerically rather than the shape/size of networks, or improve training speeds if only one network can well describe both input atomic positions and species. This opens the opportunity to develop potentials without the need to re-fit *ab initio* data when new chemical element is introduced, allowing for MLPs to *grow* rather remain limited to a fixed set of chemistries. An example would be to add chromium impurity to an existing trained FLiBe MLP. Another difficulty for general potential development is the generation of sufficient structures for training. Because molten salts contain local atomic structures characteristic of the composition and concentration of end members, it is critical to capture these local structures while requiring minimal lengthy DFT evaluations. The generation of structures purely *via* AIMD is excessively expensive considering that many structures encountered in MD may be redundant or lacking distinctive features for MLP training. Therefore, more effective sampling methods are needed to speed up this process. In addition, the transferability of our model to other compositions is tested here on FLiBe, with increasingly higher disagreement to the linearly decreasing temperature relation of the density as one departs from 33% BeF_2. Despite this, we observe that the

resulting simulations are stable and their densities comparable to the composition of the training set, implying some degree of transferability. We expect an improvement in the model with further training on other compositions, however we leave this for future work due to the expensive nature of AIMD simulations. On the other hand, active learning techniques offer the opportunity to generate data efficiently by sampling new configurations with the MLP itself, requiring DFT only when the model recognizes a poorly represented atomic configuration. The active learning approach recently presented by Zhang *et al.* addresses the bottleneck of dataset generation by exploration and labeling of configurations through DPSE-MD, cutting down the required number of DFT evaluations to 0.0044% of the total number of configurations explored [115]. Therefore, active learning can significantly improve the search for poorly represented local structures in molten salts varying with respect to temperature and composition while minimizing the needed DFT to generate the training set. Essentially, the development of machine learning potentials with improved elemental scaling and minimal DFT calculations is keystone for modeling a wide range of molten salts for rapid database development. Currently, we are working on a modification of deep learning potentials that changes how the model distinguishes elements with improved scaling in model complexity for application on molten salts and existing large databases containing materials such as ceramics and semiconductors.

Acknowledgements

This material is based upon work supported under a Department of Energy, Office of Nuclear Energy, Integrated University Program Graduate Fellowship (IUP) and Nuclear Energy University Program. Alejandro Rodriguez is financially supported by the IUP under the award number DE-NE-0000095. Stephen Lam is supported by award number DE-NE0009204. Research reported in this publication was supported in part by the NSF (award number 2030128).

Chapter 4 Million-Scale Data Integrated Deep Neural Network for Phonon Properties of Heuslers Spanning the Periodic Table[3]

[3] Rodriguez, A., Lin, C., Yang, H., Al-Fahdi, M., Shen, C., Choudhary, K., Zhao, Y., Hu, J., Cao, B., Zhang, H., and Hu, M., 2023, “Million-Scale Data Integrated Deep Neural Network for Phonon Properties of Heuslers Spanning the Periodic Table,” 9(20), Submitted to npj Computational Materials, 7/28/2022.

Abstract

Existing machine learning potentials for predicting phonon properties of crystals are typically limited on a material-to-material basis, primarily due to the exponential scaling of model complexity with the number of atomic species. We address this bottleneck with the developed Elemental Spatial Density Neural Network Force Field, namely Elemental-SDNNFF. The effectiveness and precision of our Elemental-SDNNFF approach is demonstrated on 11,866 full, half, and quaternary Heusler structures spanning 55 elements in the periodic table by prediction of complete phonon properties. Self-improvement schemes including active learning and data augmentation techniques provide an abundant 9.4 million atomic data for training. Deep insight into predicted ultralow lattice thermal conductivity (<1 W m^{-1} K^{-1}) of 774 Heusler structures are gained by p-d orbital hybridization analysis. Additionally, a recently discovered class of two-band charge-2 Weyl points, referred to as “double Weyl points,” are found in 68% and 87% of 1,662 half and 1,550 quaternary Heuslers, respectively.

4.1 Introduction

In semiconductors and insulators, phonons dominate the lattice thermal conductivity (LTC) as the quanta of atomic vibrations [49]. Because of its ubiquity, knowledge of the phonon properties such as phonon dispersions and LTC in ordered structures is enormously important in the development of innovative energy related technologies, such as energy conversion, thermal management, quantum computing, etc. Fast and accurate prediction of phonon properties is necessary for discovery of novel materials for those applications. Currently, first-principles based anharmonic lattice dynamics (ALD) method coupled with the phonon Boltzmann transport equation (BTE) is one of the most featured and accurate methods to obtain the phonon properties including LTC, which involves tedious calculations of harmonic and anharmonic interatomic force constants (IFCs) of crystalline structures relating the potential energy and atomic displacements [116]. Despite the parameter-free and predictive calculations of density functional theory (DFT), obtaining IFCs via the real-space supercell-based finite displacement method is very time and resource consuming. This situation is even worse for high-throughput computation of large number of materials [117]. Many efforts are put forth to circumvent this time-consuming nature of computing LTC with DFT, dubbed “DFT-LTC” here, to achieve high-throughput discovery of materials for target LTCs. A primary route that recently has been taken by storm is machine learning (ML). Due to their demonstrated ability to fit complex non-linear, multi-dimensional functions at orders of magnitude faster than the traditional enumeration schemes, many ML methods

have been incorporated to accelerate LTC computation. For instance, researchers have recently pursued data-driven approaches through extraction of vital information from already existing DFT-LTC data to explore previously unseen structures [118–122]. Recently, Zhu *et. al.* predicted the LTC of 92,919 materials with a 154-dimensional descriptor as the input to random forest prediction, with the three most important features for LTC prediction being the average volume per atom in the ground state, average bond length, and volume per atom [121]. Additionally, Miyazaki *et. al.* incorporated combinations of atomic radii, atomic masses, and elements from 143 half-Heusler structures into a sequence of regression models to predict the lattice parameter and thermal conductivity of unseen half-Heuslers within 1% and 4% of the DFT results, respectively [120].

While training ML models with material descriptors offers physical insights towards feature importance for LTC prediction, limitations are present when facing high-throughput. Mainly, these models are still required to generate reference LTC data to serve as the target during training. ML models such as artificial neural networks (ANNs) depend on data diversity due to their interpolative nature, i.e. they cannot perform well when provided data are outside of the training set [12,24]. As such, the data generation for a sufficiently robust model is expensive and may limit the predictions to a small subset of materials. Additionally, because these models are usually trained on one temperature designed to output a single value of LTC, they are unable to provide the plethora of information that comes with DFT-LTC calculations [118–121]. Outputs such as

phonon dispersion, scattering rates, temperature-dependent LTC, and off-diagonal thermal conductivity values in the LTC tensor are inaccessible, all of which are standard outputs from phonon calculators [47,48,123,124].

To circumvent these issues, the LTC may be approached from a lower-level, more specifically through the atomic forces, which are the fundamental input and starting point of the DFT-LTC procedure. Approaching BTE solvers with already computed atomic forces from ML maintains the rich output of information that comes standardly from phonon calculators. More importantly, training for the atomic forces has the potential to reduce the costly demand for training set generation. For traditional ML models, many DFT calculations are required for one LTC value, serving as a single datapoint for training. In contrast, training on atomic descriptors is advantageous in terms of data abundance per DFT run because each simulation provides (3*N*+1) data corresponding to *N* atoms worth of force vectors and one total energy. Namely, machine learning potentials (MLPs) implicitly capture the electronic-level features from DFT by representation of the potential energy surfaces as functions of the atomic nuclei positions [12]. Due to the purely mathematical nature of ML models, the accuracy of MLPs strongly depends on the description of the atomic environment surrounding central atoms to capture the appropriate physics [125]. Many studies have shown excellent representation of DFT-level energetics and realistic property prediction with MLPs such as the High Dimensional Neural Network Potential (HDNNP) [13], Deep Potential Molecular Dynamics (DeePMD) [113], and Gaussian Approximation Potential (GAP) with SOAP descriptors [126]. In the context of

phonon property prediction, several studies have used MLPs as the force calculator [22,23]. Typically, these MLPs own a root mean square error (RMSE) of the force predictions within 10 – 100 meV/Å with approximately 10^3 faster evaluation time compared to DFT.

Undoubtedly, the robustness of the MLP has the potential to mitigate the current speed-related bottlenecks in the LTC workflow. However, to-date the majority of studies using MLPs share a common denominator in that the models are limited to a material-to-material basis. This is primarily due to the exponential scaling of model parameters with the number of atomic species or elements (N_{elem}). For example, the HDNNP requires N_{elem} element-specific networks each containing approximately N_{elem} radial and $N_{elem}(N_{elem}+1)$ angular symmetry functions [66]. When faced with data containing elements spanning the periodic table, the training efficiency and evaluation time is reduced significantly due to the $\sim N_{elem}^2$ scaling of the input descriptors. Additionally, training of each element-specific network requires central atoms dedicated only to said element, meaning that little to no knowledge of atomic environments from other central atom species are shared. In general, recent MLPs represent atomic positions numerically, while the atomic elements depend on specific sub-models and/or order of the input descriptors, in-turn diminishing prediction quality with ten or more elements. Overall, independent elemental scaling and centralized ML training are two major factors necessary for evaluation of theoretical materials databases containing a plethora of structures and atomic species that would otherwise be too difficult to handle with modern MLPs. Computing forces across

many atomic environments is especially challenging for high-throughput LTC considering the notoriously strict force accuracy requirements for the IFCs and the resulting LTC [96,127].

In this work, we developed accurate force calculators called Elemental Spatial Density Neural Network Force Field (Elemental-SDNNFF) including high scope of transferability between atomic structures and elements. The Elemental-SDNNFF shows an unprecedented < 10 meV/Å force error for the atomic forces covering 11,866 structures including 55 elements, effectively spanning the periodic table as detailed in section 4.2. To easily access millions of data from costly DFT calculations, training on atomic forces allows an *N*-fold increase granting the current model abundant high-resolution force information, whereby *N* is the number of atoms per supercell. This is distinct from existing MLPs especially those with total energy as the output: although the addition of other properties like atomic forces and virial terms is standard, they involve a summation over all other properties providing a *single* training point per DFT run [13,35]. Here, we further the existing data size with rotational covariance and data augmentation techniques, allowing for an average three-fold increase in the data size as explained in sections 4.3 and 4.4. We also incorporate active learning techniques to generate data with little to no human intervention with the "query by committee" method, which is based on the uncertainty of atomic forces between several models as detailed in section 4.5. Sections 4.6-4.9 provide information about structure generation, data preparation, and model parameters and details. In sections 4.11-4.16, we demonstrate these methods by training and

applying the elemental-SDNNFF to a database containing quaternary (ABCD), half (ABC), and full (ABC_2) Heusler structures, all of which are trending in research due to their capacity for high thermoelectric performance [120,128,129]. Because the atomic forces are predicted directly, our results show promise for prediction of accurate high-throughput full phonon properties, such as phonon dispersions for thermodynamic stability and LTC calculations, at the fraction of the computational cost of traditional DFT-LTC. Such workflow may be easily *extended to broader types of crystals* such as noncubic structures with an arbitrary number of elements *spanning the periodic table*. Moreover, the concept of topological quantum states has recently been introduced to phonon systems [130]. Among them, Weyl semimetals in three dimensions are found in realistic materials [131,132]. Weyl points (WPs) are crossing points of two phonon bands, which can be described by the chiral Weyl equation as detailed in section 4.10. The band degeneracy is robust against small perturbations, meaning the position of a WP can move under small perturbations, but it will not disappear unless annihilated with an opposite WP. Topological protected Fermi arc surface states connecting the projections of a pair of WPs can be found on the surface Brillouin zone [133]. Before this work, the predictions of WPs were based on DFT-derived phonon dispersions. Our elemental-SDNNFF spanning the periodic table offers vast opportunities for exploring large number of candidate Weyl semimetals in a more efficient way.

4.2 Model Descriptors

The SDNNFF was originally inspired by the HDNNP wherein each atomic descriptor is a summation of atomic contributions in radially-dependent functions [16]. However, unlike in MLPs, the SDNNFF is designed to only model the atomic forces without the total energy, a so-called neural network force field (NNFF). In situations where the total energy is not required, an NNFF provides two major advantages: a) the resolution of training on individual atomic force vectors in NNFFs, rather than a function of the total energy plus the *summation* of force components in MLPs, significantly augments (generally by two orders of magnitude) the available training data from t to $N \times t$, where t is the number of supercells evaluated by DFT and N is the number of atoms in the supercell. Each DFT run, therefore, yields N data for training, and b) prediction of the force vector directly eliminates the need to calculate the derivative of the total energy with respect to network inputs, or $\boldsymbol{F}_i = -\nabla_i E$ providing computational cost and time savings in training and evaluation. The former N data per supercell is the greatest motivator for SDNNFF development due to the improved yield in training data per costly DFT run. Although traditional MLPs also can take advantage of the force information, they are involved as a summation between DFT and prediction forces in the loss function for compatibility with the single energy value, effectively reducing the resolution of atomic forces and the overall number of training data [13,35]. This can be evidently seen from the RMSE for forces: <10 meV/Å for our Elemental-SDNNFF as compared to the several tens and even hundreds meV/Å for previous MLPs. Furthermore, since the IFCs for LTC

calculation require only the displaced positions and forces of atoms, the absence of the total energy in this application is not an issue.

In our previous development of SDNNFF, the atomic environment is represented by a functional mapping of the 3D space rather than the polar space [96]. Following the previous model, the current solution for descriptor development including both atomic positions and elements requires only a single network scaling independently with respect to the available species in the training set. Following the same definitions as provided in Chapter 3, the modified SDNNFF model, dubbed the Elemental-SDNNFF, uses the following descriptors as input

$$\varphi_{\boldsymbol{\alpha}}^{0} = \sum_{n=1}^{N} \begin{cases} \frac{1}{2} f_c(|\boldsymbol{R}_n|) * \left(\cos\left(\frac{\pi}{\sqrt{3}L} R_n^{\boldsymbol{\alpha}} \right) + 1 \right) & \text{if } R_n^{\boldsymbol{\alpha}} < D \times L \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

$$\varphi_{\boldsymbol{\alpha}}^{1} = \sum_{n=1}^{N} \begin{cases} \frac{W_n}{2} f_c(|\boldsymbol{R}_n|) * \left(\cos\left(\frac{\pi}{\sqrt{3}L} R_n^{\boldsymbol{\alpha}} \right) + 1 \right) & \text{if } R_n^{\boldsymbol{\alpha}} < D \times L \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

$$\boldsymbol{\varphi} = \text{concat}\left(W_{central}, \text{concat}\left(\varphi_{\boldsymbol{\alpha}}^{0} \text{ for all } \boldsymbol{\alpha} \right), \text{concat}\left(\varphi_{\boldsymbol{\alpha}}^{\mathbf{1}} \text{ for all } \boldsymbol{\alpha} \right) \right) \tag{4.3}$$

where $\varphi_{\boldsymbol{\alpha}}^{0}$ is similar to the previous SDNNFF descriptor representing purely the spatial distribution of atoms, $\varphi_{\boldsymbol{\alpha}}^{1}$ is the spatial-elemental descriptor, W_n is the atomic number of neighboring atom n, and $W_{central}$ is the atomic number of the central atom. Additionally, the cutoff function $f_c\left(\left|\overrightarrow{R_n}\right|\right)$ is defined as

$$f_c(|\boldsymbol{R}_n|) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{\pi|\boldsymbol{R}_n|}{R_c}\right)+1\right] & \text{for } |\boldsymbol{R}_n| \leq R_c \\ 0 \text{ for } |\boldsymbol{R}_n| > R_c \end{cases} \tag{4.4}$$

and was added to represent the decaying influence of atom n from the central atom, observably improving the force accuracy. Also, the only difference between φ_{α}^{0} and φ_{α}^{1} is the factor W_n where the density function is multiplied by the corresponding atomic number of neighboring atom n. As a result, the additional cost from φ_{α}^{1} is minimal since the already computed values from φ_{α}^{0} are simply multiplied with the corresponding atomic weights. Furthermore, $\boldsymbol{\varphi}$ is the finalized descriptor vector in which the central atom atomic number, the spatial descriptor vector, and the spatial-elemental descriptor vector are all concatenated in 1D. The basic idea of the added descriptor is to simultaneously capture the previously accurate spatial mapping of neighbors in addition to the influence of atomic elements on the signals measured at the same grid points. Two advantages arise from the elemental-SDNNFF descriptor: (1) The summation of weighted density functions in φ_{α}^{1} eliminates the need for designated slots in the descriptor vector for each element and removes the scaling of input size with respect to number of elements, and (2) By providing $W_{central}$ in $\boldsymbol{\varphi}$, the network can distinguish central atoms whereby individual element-specific SDNNFFs are not required. The result is a singular NNFF capable of modeling atomic systems spanning the periodic table without sacrificing network efficiency, demanding only one network for training with a fixed two times plus one inputs as the previous SDNNFF model.

4.3 Rotational Covariance

As mentioned previously, the SDNNFF is constructed by a 3D mapping of space corresponding to the 3D forces as the output of the network. In the original version of the SDNNFF, the reference coordinate system was constructed by the input coordinates to the DFT system from the structure file containing atomic positions and Bravais lattice vectors. The grid point positions and consequentially the descriptor was dependent on the coordinate system of the reference DFT data the grid is always built along the reference *xyz* directions. As a result, a rotation of the atoms in the system can yield dramatic changes to the input descriptor, and given the purely mathematical nature of neural networks, the resulting forces are likely to mismatch with those prior to the same rotation. Thus, it is beneficial to design an NNFF with rotational covariance. The advantage of rotational covariance is the capacity to model infinitely many possible rotations with fewer equivalent representations, reducing the redundancy in training similar but rotated atomic systems. Rotational covariance also helps reduce the number of DFT configurations needed for force accuracy convergence since redundant atomic neighborhoods of existing but rotated systems are already considered.

One simple method to incorporate rotational covariance is to construct a rotation matrix based on the positions of the two radially nearest neighbors [91]. If the relative position of the first closest neighbor is $\boldsymbol{R}_1$ and that of the second closest neighbor is $\boldsymbol{R}_2$ then

$$\boldsymbol{e}_1 = e(\boldsymbol{R}_1) \tag{4.5}$$

$$\boldsymbol{e}_2 = e(\boldsymbol{R}_2 - (\boldsymbol{R}_2 \cdot \boldsymbol{e}_1)\boldsymbol{e}_1) \tag{4.6}$$

$$\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2 \tag{4.7}$$

where $e(\boldsymbol{R})$ represents the normalized vector or $e(\boldsymbol{R}) = \boldsymbol{R} / |\boldsymbol{R}|$, and $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ are the unit vectors forming the orthogonal axes for the rotated system. The rotation matrix **M** and the transformation of atom n from the unrotated atomic coordinate $\boldsymbol{R}_n = (x_n, y_n, z_n)$ to the rotated coordinate $\boldsymbol{R}'_n = (x'_n, y'_n, z'_n)$ is thus

$$\mathbf{M}(\boldsymbol{R}_1, \boldsymbol{R}_2) = \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{e}_3 \end{bmatrix}^{\mathrm{T}} \tag{4.8}$$

$$\left(x'_n, y'_n, z'_n\right) = \left(x_n, y_n, z_n\right) \cdot \mathbf{M}(\boldsymbol{R}_1, \boldsymbol{R}_2) \tag{4.9}$$

For two or more systems that differ by an arbitrary rotation, since the rotation matrix **M** is constructed on the rule that $\boldsymbol{R}_1$ is always the first closest neighbor and $\boldsymbol{R}_2$ is always the second closest neighbor, then the same rotated atomic coordinates $\boldsymbol{R}'_n$ are always obtained regardless of the rotation performed on the system. In practice, one problem that is encountered especially in ordered structures such as crystals is that the possibility of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are set exactly 180° apart with respect to the central atom. This makes $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2 = 0$ thus resulting in the absence of the third axis and the transformation cannot be performed. As such, the above rule is modified to make $\boldsymbol{R}_2$ such that a) the atom is coordinated, i.e. belongs to the set of atoms closest to the central atom in equilibrium and b) the atom forms the smallest angle with the central atom and $\boldsymbol{R}_1$. For the cases here, it so happens this rule prevents the selection of any two

atoms that are 180° with respect to the central atom. Figure 4.1 shows an example coordinated neighborhood containing a central (green) atom. The blue atom is the closest to the central atom, and the red atom forms the smallest angle out of all nearest neighbors. The brown atoms form angles greater than that of the red atom and are neglected. Then, both the blue and red atom positions are then used for generating the unit vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$. From these unit vectors, the black lines show the axes forming the rotation matrix.

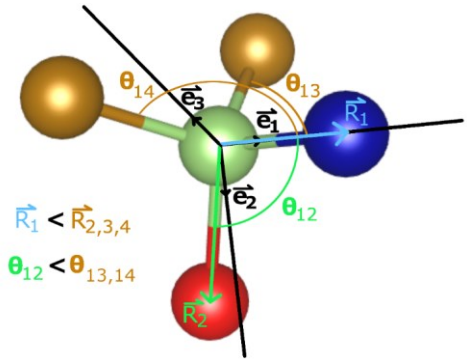

Figure 4.1 Schematic showing the nearest neighbor selection rules for generation of the rotation matrix.

Additionally, provided rotation of the local atomic environment by matrix **M**, the input descriptor is changed in comparison to that of the unrotated case. As a result, the problem arises that the rotated descriptor no longer lies in the original coordinate system of the DFT forces, and the training on these forces requires additional treatment. Thus, the approach here is to modify the architecture of the neural network model to train on rotationally covariant inputs and yield forces in the same crystalline coordinate system. Figure 4.2 provides a schematic example of the neural network model for the current SDNNFF training.

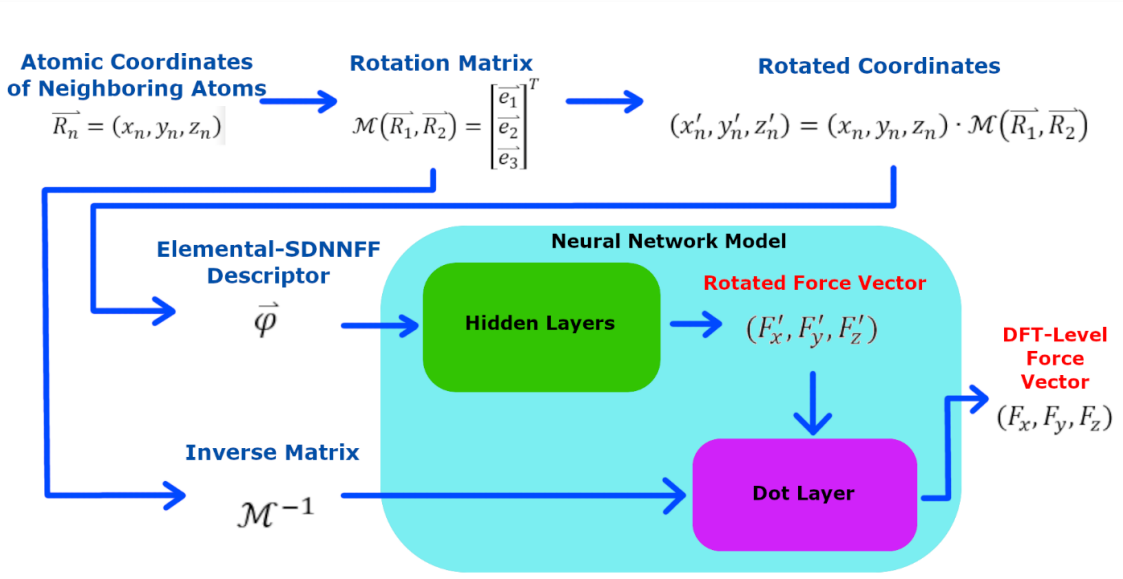

Figure 4.2 Schematic of the neural network architecture considering rotational covariance.

First, the rotationally covariant elemental SDNNFF input is generated from existing DFT data. Simultaneously, the inverse of the rotation matrix $\mathbf{M}^{-1}$ is saved and serves as an input to the network model. The generation of **M** is necessary for the descriptor vector and taking its inverse is relatively insignificant in terms of cost. Secondly, the input descriptor passes through several hidden layers. Thirdly, after the hidden layers, the number of nodes is deliberately set to three to represent a 'pseudo-force' vector described by the rotationally covariant nature of the input. During application, this force vector alone cannot represent the DFT force, so an additional dot layer is added. The dot layer multiplies the vector and the matrix $\mathbf{M}^{-1}$, effectively converting the vector back into the cartesian space of the atomic system. Finally, this result yields the DFT-level forces from which the training program computes the loss function with respect to the reference DFT

forces and performs back-propagation. Because the trainable parameters of the network lie between the rotationally covariant input descriptor and the dot layer, the model is trained on rotationally covariant information and is applicable to systems regardless of rotation **M**.

4.4 Data Augmentation

One of the objectives of this work is to minimize the number of DFT calculations for dataset generation. Following the existing $N \times t$ scaling of the dataset, gathering as much information as possible from each DFT run is imperative to improve the speed/cost ratio of NNFF training and evaluation. If the number of DFT calculations for network generation is close to or exceeding that required for LTC calculations for all materials in the data pool, then the NNFF quickly loses its novelty; the time for dataset generation and training could had been spent directly on LTC instead. As such, furthering the $N \times t$ scaling is therefore a critical aspect for the Elemental-SDNNFF. Here, the selection rule for rotational covariance is discussed in which two neighboring atoms relative to the central atom are selected: the first atom is the closest atom, and the second atom is a) coordinated with the central atom and b) forms the smallest angle with the first atom and the central atom. In this case, when provided a crystal with displaced atoms, the choice in the first and second atoms may vary greatly despite the similarity of atomic environments. This may artificially create gaps in knowledge due to the seemingly sporadic nature of the displaced atoms and the resulting rotation matrix. A simple way to fill these gaps in the data is to take several candidates for the second atom, i.e., those that are coordinated with the

central atom but own similar angles as that with the smallest formed angle. These candidates are then used for rotation matrix generation for the same atomic environment and are included in the training set. Figure 4.3 shows an example of data augmentation performed on a single atom in ABCD structure provided a finite cutoff up to the first neighbor for clarity. The bottom right of Figure 4.3 shows the closest atom as overlayed with a blue circle and similarly the second atom with green. In addition to the original scheme, similar atomic environments provided the central and selected blue atoms may yield two more possibilities for the second atom. As a result, a 3× increase in data for training is expected in atomic environments for ABCD structure, as also observed in ABC and ABC_2 structures in general and the dataset is augmented equally throughout.

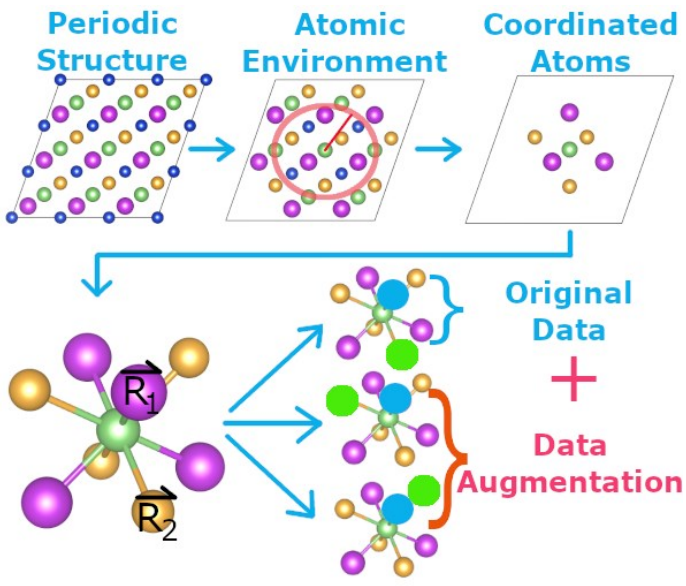

Figure 4.3 Data augmentation of a single central atom performed on a displaced quaternary Heusler structure.

4.5 Active Learning

An important aspect of any ML process is to consider the quality of the dataset used for training. Ideally, the training data should contain sufficiently diverse data to cover all possible features within the domain of the application. However, judging the so-called quality and diversity of atomic force fields based on the simultaneous positions of atomic natures is not so trivial. Inspired by the work of Zhang *et. al.*, active learning is incorporated for Elemental-SDNNFF dataset generation and self-improvement [115]. Specifically, the ‘query by committee’ method is a form of active learning wherein several identical models are trained in parallel on the same dataset but own different initialized weights. After training, these models collectively evaluate a pool of structures, and a comparison of the predicted forces is performed. If the variance in the resulting forces is low for a particular structure, then the associated atomic configurations contain features that are well considered in the training set. On the other hand, due to the interpolative nature of neural networks, a high variance implies that the atomic configuration owns features outside of the dataset and the structure is pooled as a candidate for retraining. With this pool, DFT is performed, and the new data is retrained for all the models in the committee, and the loop is repeated. Here, several Elemental-SDNNFF networks are trained on the initial dataset and serve as the committee. After training, we evaluate the atomic forces on a pre-generated set of displaced supercells. Like the work by Zhang *et. al.,* an indicator is computed for each atomic configuration serving as the uncertainty observed by the committee [115].

The per-atom indicator in this work is defined as

$$\varepsilon_i = \sqrt{\left\langle \left\| f_i^m - \bar{f}_i \right\|^2 \right\rangle}, \bar{f}_i = \left\langle f_i^m \right\rangle \tag{4.10}$$

where ε_i is the indicator for atom i, f_i^m is the predicted force by model m, and $\bar{f}_i$ is the average force across all models. Essentially, ε_i is designed to *indicate* whether a particular atomic force is well represented by the models. Low ε_i equates to low variability of the predicted forces across all models and the atomic configuration is well-considered, whereas increasing ε_i depicts a higher degree of uncertainty and the models require further improvement. Note, a low ε_i does not always guarantee low errors relative to DFT, since configurations close to those in the training set may own high sensitivities to slight shifts in positions from DFT. However, a high ε_i is strongly indicative of a poorly extrapolated configuration regardless of the DFT result and should be considered as a candidate for model retraining.

Candidates for subsequent rounds of DFT evaluation and training are filtered through several steps. As seen in Figure 4.4, the first step is to perform structure optimization with DFT followed by artificial generation of several displaced supercells for all available structures. The second step is to pass these supercells through the committee of models and evaluate all atomic configurations, returning ε_i for each atom. After force evaluation with the committee models, the third step is to select candidate supercells containing large variances in atomic forces. Specifically, structures not satisfying $\varepsilon_i^L < \varepsilon_i < \varepsilon_i^U$ for all atoms are filtered out, where ε_i^L is the lower bound and ε_i^U is the upper

bound for the indicator. The choice of the lower bound ε_i^L should depend on the order of the force error in the committee of models. A lower bound identical to the force error is problematic since candidates may be chosen based on the intrinsic error of the network. So, a value sufficiently higher than the force error for the lower bound should be selected. The choice of the upper bound ε_i^U is somewhat arbitrary, but it should act as a filter for unreasonable atomic configurations, e.g., due to an unrealistic structural configuration. Then, the remaining supercells are sorted based on the maximum value of ε_i over all atoms i, or $\max(\varepsilon_i)$, to prioritize configurations containing the highest discrepancy of the predicted forces. From this list of supercells, a single supercell corresponding with the maximum of $\max(\varepsilon_i)$ is chosen as the candidate among all supercells for its associated unique structure. As a result, the final list of candidates will contain at most one supercell from each structure, and zero supercells from structures not satisfying $\varepsilon_i^L < \max(\varepsilon_i) < \varepsilon_i^U$. Although several supercells may exist with high $\max(\varepsilon_i)$ for a corresponding structure, only the top $\max(\varepsilon_i)$ is picked to guarantee variety in the final list of candidates. The fourth step is pooling, whereby the pool is sorted by the number of atoms satisfying $\varepsilon_i^L < \varepsilon_i < \varepsilon_i^U$. Because each DFT run provides several atomic forces, maximizing the number of atoms with $\varepsilon_i^L < \varepsilon_i < \varepsilon_i^U$ also maximizes the ratio of previously unseen datapoints to new DFT calculations which is of interest. Finally, the top few supercells from the final pool are chosen for DFT evaluation and retraining in the committee of models. Note that although the choice in number of supercells from the final pool for DFT calculation was somewhat arbitrary (in the range of 100 – 1,000 supercells), this was primarily

dictated by the availability of computational resources. Additionally, structures not retrained into the network are expected to resurface as candidates in subsequent active learning iterations.

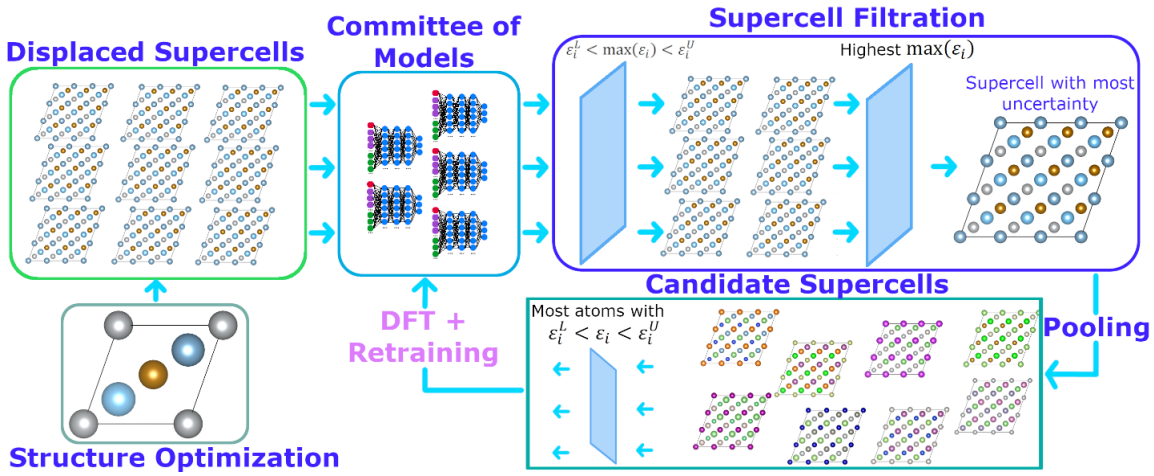

Figure 4.4 Active learning loop for model improvement with the query by committee method.

4.6 Structure Generation

Initially, a small pool of ABCD structures serve as the training set for baseline model training. First, the primitive unit cells for the structures are relaxed with DFT using the same parameters for consistency. Then, the supercells are generated using PHONONPY by duplicating the cell by 3×3×3 and displacing all 108 atoms in random directions by a fixed 0.03 Å distance [47]. Specifically, 1,000 structures with 5 displaced supercells each, 500 structures with 6 supercells each, and 1,500 equilibrium supercells are generated, yielding 9500 structures or 1.026×10^6 unique atomic environments. Finally, these structures are passed through DFT to generate the reference data for initial training.

Subsequent additions of data are primarily done through active learning, with some data from DFT-LTC calculations obtained by comparisons between iterations of the Elemental-SDNNFF model and DFT. Note, the ABC and ABC_2 structures used in this work are also relaxed and duplicated to 3×3×3 supercells containing 81 and 108 atoms, respectively. Figure 4.5 shows the pipeline of the data involved in this work. Specifically, after structure optimization, the remaining 11,866 structures are used to generate randomly displaced supercells in two separate groups, one for active learning and the other for LTC calculation. This means that the supercells in active learning are not the same as those in the LTC group to guarantee little to no bias when evaluating the LTC of some structures from the 11,866 set. After training with the active learning procedure, the LTC is calculated for a separate set of 1,298 structures not included in the training. Note, 50 supercells per structure are also generated for LTC calculation of the testing set.

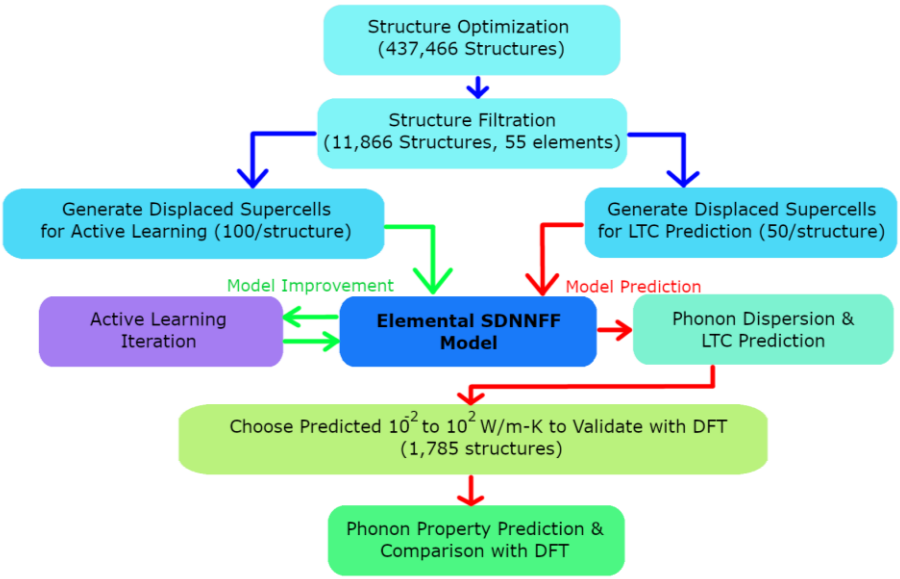

Figure 4.5 Overall data pipeline for model development and evaluation.

4.7 Data Preparation

To develop a pool of stable structures for this study, several filtration steps were done on a large pool ABCD, ABC, and ABC_2 Heusler structures with space group number 216, 216, and 225, respectively. The initial configuration of these structures were borrowed from the Open Quantum Materials Database (OQMD) lacking LTC data [134] and were then reoptimized by the Vienna Ab initio Simulation Package (VASP) [135–137] using our own parameters. As seen in Figure 4.6.a, a pipeline is performed here for full, half, and quaternary Heusler structures. Numbers on the right of the pipeline represent the number of ABC2 (red), ABC (green), and ABCD (blue) structures after passing through the adjacent filter. The final filter has a slightly reduced structure count from the removal of <0.1 and >200 $Wm^{-1}K^{-1}$ structures, which are assumed to be outliers.

Figure 4.6.b shows the structures studied here with the corresponding number of structures from OQMD. Figure 4.6.c provides the total number of elements corresponding to central atoms included in training the Elemental-SDNNFF. Elements without color are not included in this study. Here, the first step is to filter out structures containing lanthanide and actinide elements to limit the number of structures in this study for computational reasons, although future studies including these elements is certainly of consideration. Then, the structures are filtered by the formation energy after structure optimization with DFT, where lower formation energies have higher tendency to be stable. The formation energy is quick to compute using DFT requiring only the primitive cells comprising of 3 – 4 atoms. Finally, the energy above hull (E_{hull}) provides the ground state stability of partial compounds with respect to all possible linear combinations of phases present in the compound phase diagram, which is also not time consuming. The low E_{hull} value has higher probability to yield thermodynamically stable structures, i.e., all positive frequencies in the phonon dispersion. The final pool in this study holds 2,377 quaternary Heusler (ABCD), 2,660 half-Heusler (ABC), and 6,829 full-Heusler (ABC_2) structures (totaling 11,866) from which the Elemental-SDNNFF model training and prediction of atomic forces, phonon dispersions, and LTC values are performed.

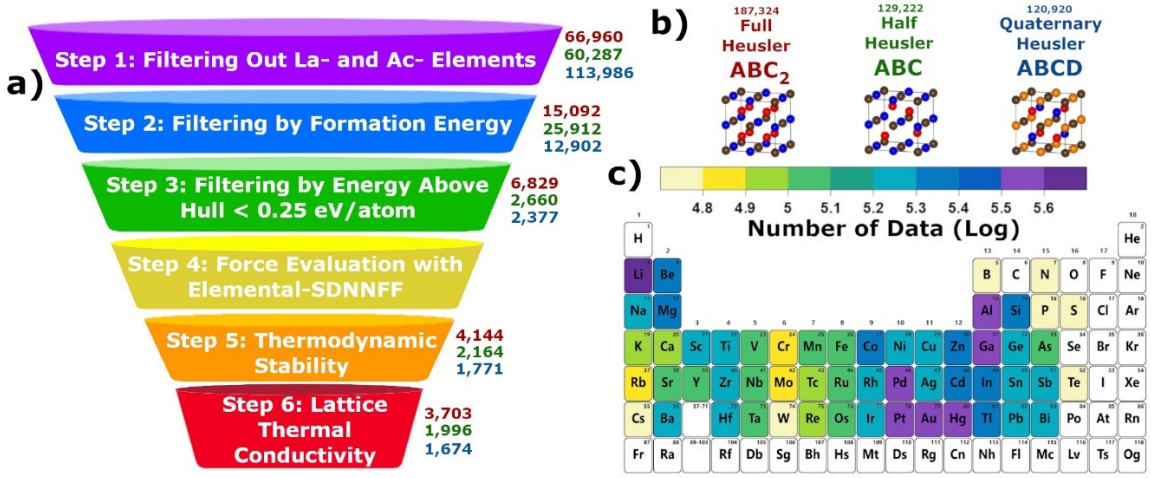

Figure 4.6 Summary of the dataset generation in this work.

4.8 Model Details

The neural network models receiving Elemental-SDNNFF inputs are constructed and trained on Python using Tensorflow 2.8.0 [92]. Specifically, the networks initially contained 3 hidden layers and 500 nodes with a default mini-batch size of 2048. The He uniform initializer was used to generate the initial weights and biases of the network [44], the Adam algorithm was used as the optimizer for updating the weights with a learning rate of 0.0001 [45], and the exponential linear unit (ELU) was implemented as the activation function of choice [46].

The cutoff for atomic interactions was initially chosen as 6 Å and the resolution was set to k = 10 providing 721 as the input descriptor length. Here, the 3rd-order IFCs are truncated at the third neighbors for computing LTC of all structures, meaning that the displacements of atoms beyond this cutoff are neglected. Since truncation of the interactions beyond the third neighbors

typically yields converged LTC, and because the third neighbors in the pool of ABCD do not exceed 6 Å, the choice was sufficient for the time being. However, subsequent inclusion of full and half Heusler structures owning smaller atomic densities contained third neighbors beyond 6 Å. Specifically, the (average third neighbor distance in Å, max third neighbor distance in Å) for each ABC2, ABC, and ABCD in the dataset are (5.746, 7.427), (6.971, 11.474), and (5.305, 6.754), respectively. Nonetheless, for generation of the dataset from active learning, we maintain the 6 Å cutoff anticipating improvements in the phonon frequencies and LTC from future retraining of networks with higher cutoffs. After several rounds of active learning, 7 Å and 8 Å networks with k = 12 (1,345 input descriptors) and seven hidden layers are trained from the ground-up with the overall dataset to test the sensitivity of the phonon frequencies and LTC with respect to force cutoff. The details of network performance with modified cutoff are described in section 4.11.

4.9 Phonon Calculation Details

Calculation of the 2nd and 3rd-order IFCs using predicted Elemental-SDNNFF forces were performed using compressive sensing lattice dynamics (CSLD) method [138]. In short, CSLD extracts the IFCs from the Taylor-expanded interatomic forces in terms of atomic displacements *via* least squares or more advanced compressive sensing, using a user-specified number of supercells containing atoms that are either randomly displaced or generated with molecular dynamics. The advantage of this method is the significantly lowered requirement of supercell number for IFCs, reducing the needed DFT for converged LTC. For

example, for quaternary (ABCD, space group number: 225) Heusler, traditional finite displacement method would require at magnitude 10^2 structures to get the 3rd order interatomic force constants even if only third nearest neighbor is considered, while only 5-12 randomly displaced structures are needed from CSLD. By rule of thumb, the number of randomly displaced supercells per structure with DFT is 6 for ABCD, 12 for ABC, and 5 for ABC_2 structures. On the other hand, we found that 50 supercells were sufficient for LTC convergence using SDNNFF. Due to the stochastic nature of CSLD, more supercells are needed to compensate for the error of the force predictions from the model. Nonetheless, the speed of the SDNNFF drew no concern for the calculation. The interactions for the 3rd-order IFCs were truncated at the third nearest neighbors for all structures. All LTC calculations were performed by ShengBTE package [139] at a temperature of 300K with a 24×24×24 q-point sampling grid. For verification, 443 ABCD, 506 ABC, and 349 ABC_2 totaling 1,298 structures are set aside for DFT-LTC calculation for comparison with the Elemental-SDNNFF. These structures are not trained in the model and are selected from previously unseen OQMD structures by Elemental-SDNNFF predictions to guarantee coverage for a wide range of LTC, including low (< 1 W/mK), medium (1 W/mK to 40 W/mK), and high (> 40 W/mK) LTC structures.

4.10 Weyl Points Searching

The existence of WPs requires broken of either time reversal symmetry or inversion symmetry [130]. Due to the lack of time-reversal-breaking mechanism, only non-centrosymmetric materials can have WPs. Here the candidates are

1,662 ABC and 1,550 ABCD Heusler structures with non-centrosymmetric space group number 216 and phonon dispersions predicted by our Elemental-SDNNFF model. For each material, we first search for two-fold degenerate points, then determine the Chern numbers of each degenerate node. For the searching of gap-closing points, we use a 10×10×10 Γ center mesh as starting points. Then, a Limited-memory Broyden Fletcher Goldfarb Shanno Bound (L-BFGS-B) optimization procedure is applied to find the local minimum of the gap between two adjacent bands. To make sure the nodal point is isolated, i.e., it is not a point on a nodal line or surface, the gap on a surrounding surface is checked. After collecting the nodal points, they are transformed into the irreducible Brillouin zone (IBZ) by applying symmetry operations. Notice that the Chern numbers of two equivalent k-points are associated by the determinant of their rotation matrix. We only consider nodal points in the IBZ. Finally, we compute the Chern numbers of each nodal point using the Wannier charge center evolution approach [140,141]. Those with nonzero Chern numbers are identified as WPs.

4.11 Active Learning Results

For the active learning procedure, 2,377 ABCD, 2,660 ABC, and 6,829 ABC_2 structures totaling 11,866 serve as the untrained samples originating from the filtration performed in section 4.7. In each unique structure, 100 randomly displaced 3×3×3 supercells are generated with 0.03 Å displacement distance. Equivalently, the number of initially untrained atomic environments is 2.57×10^7, 2.15×10^7, and 7.37×10^7 for ABCD, ABC, and ABC_2, respectively. Because the error of the trained networks range between 6 – 10 meV/Å, the lower-end of the

indicator limit is set well above this error at ε_i^L = 50 meV/Å, whereas the upper limit is set to ε_i^U = 200 meV/Å. After 21 iterations of retraining, 15 involving active learning iterations, the final dataset grew to 3.12×10^6 unique atomic configurations and is increased by a factor of 3 to 9.36×10^6 after data augmentation. When compared with the available 32,137 supercells for MLP training, this is a major leap in the dataset size and is inherently due to the $N\times t$ scaling from training on atomic forces. From Figure 4.7, only 18.8% of the data is from active learning, whereas the remaining 32.9% and 48.3% are from the initial dataset and DFT-LTC data, respectively. Overall, 55 elements in Figure 4.8 are included in this dataset and trained into the model, which is unprecedentedly large for modern MLPs.

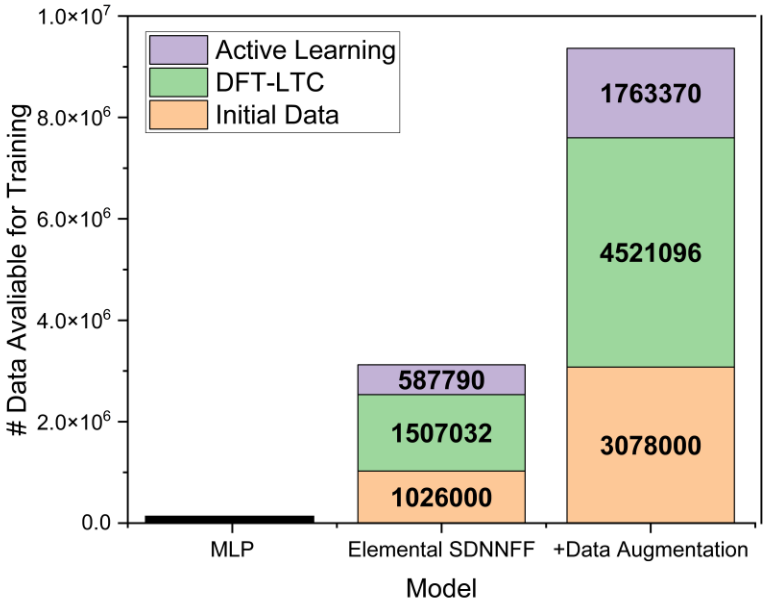

Figure 4.7 The dataset size generated from various methods.

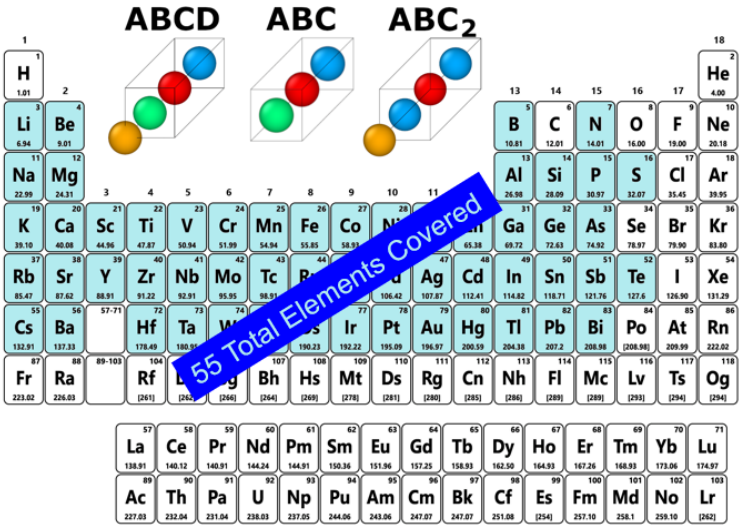

Figure 4.8 Elements included in the model training.

As seen from the learning curves in Figure 4.9, a new model is trained with 5 hidden layers on the 15th round (green dots) yielding improved RMSE when compared to the 3-layer case (orange dots), suggesting improved flexibility of the model with increased weights. Notice, new models are trained for longer epochs since they hold no memory of previous training iterations. At the 21st round, new models are again generated this time with 7 layers, higher resolution (K=12), and higher cutoffs (solid green for 7 Å and solid blue for 8 Å) yielding even lower RMSE than the R_c = 6 Å, k = 10, number of layers = 5 model (denoted as 6 Å - 10k - 5L in the figure, shown as purple dots). The most likely cause is due to the increase in layers as evidenced by the earlier improvement from 3 to 5 layers. Increased cutoff also shows slight improvement from the 7 Å to 8 Å model which may help capture forces in low density structures where the interactions are expanded. Although the change in resolution is not tested here,

most likely this also provides improvement due to the increased sensitivity of atomic displacements in the input descriptors.

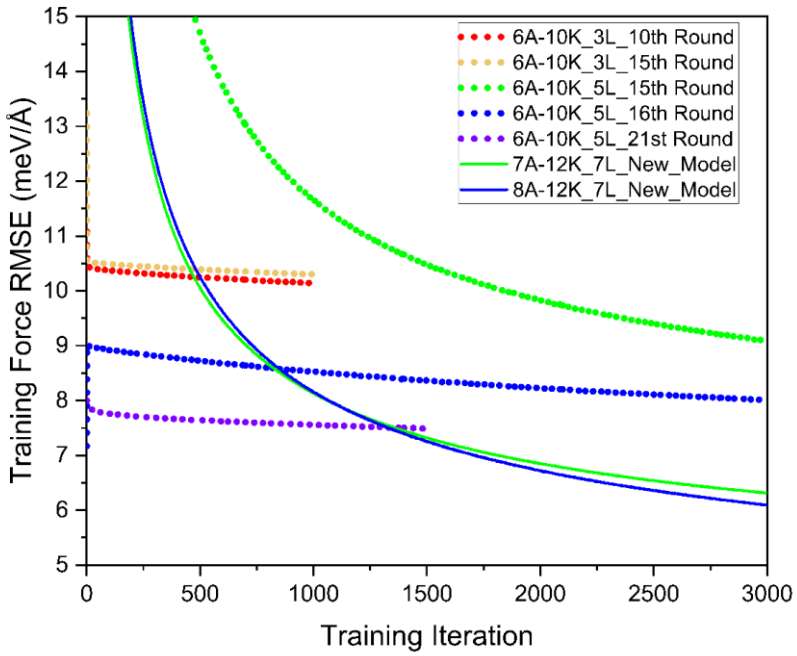

Figure 4.9 The learning curves involved with training the model over 21 rounds of active learning.

Figure 4.10 summarizes the progression of the active learning as a function of iterations broken down into ABCD, ABC, and ABC_2 structures. Here, active learning is performed on all structures whereby certain rounds are dedicated toward one or more specific structures. For instance, rounds 1-5 involve training of ABCD, whereby rounds 14 and 15 involve all three structures. Also recall that these plots only involve active learning rounds, i.e., they do not show the remaining 6 DFT-LTC rounds. In the left plot of Figure 4.10, a consistent decrease in the atoms above the lower threshold for all three

structures are observed and is expected with the subsequent new training data. However, the right of Figure 4.10 shows the efficiency from active learning as defined in Equation 4.10 and provides further insight. Specifically, the initial percent efficiencies for each ABCD, ABC, and ABC_2 are 1079.6%, 4802.1%, and 20796.9%, respectively. The initial ABCD efficiency is the lowest most likely due to the initial model training on purely ABCD DFT-LTC data and had less to learn from the displaced structures. Additionally, ABC structure shows a 4.45× increase in starting efficiency when compared with ABCD and agrees with the absence of ABC data up until round 6. Interestingly, despite sharing similar structure with ABCD, ABC_2 shows a significantly higher initial efficiency than ABCD and ABC_2 which is most likely due to the significantly higher pool of unseen 73.7 million ABC_2 atoms. Nonetheless, both ABCD and ABC_2 structures eventually converge below the 100% line, where active learning is no longer providing more knowledge than produced data. Due to time limitations, further active learning iterations necessary to achieve below 100% are not performed for ABC. Additionally, ABCD experiences little oscillatory behavior whereas ABC_2 and ABC do not experience constant downward trends. The most probable cause is the injection of DFT-LTC data after the first five active learning rounds of ABCD. Specifically, DFT-LTC data is trained after rounds 3, 8-11, and 13-15 of active learning, matching regions of little change in atoms above the threshold (left figure) and sharp increases of efficiency (right figure) for ABC and ABC_2. Nonetheless, we observe a significant improvement in the atomic force evaluation for 2.57×10^7, 2.15×10^7, and 7.37×10^7 ABCD, ABC, and ABC_2 atoms

with only 6.7×10^4, 1.08×10^5, and 4.13×10^5 additional atoms trained from active learning over all 11,866 structures in the dataset. By allowing models to choose unknown data, active learning allows growth in models with little human interference in the structure selection process.

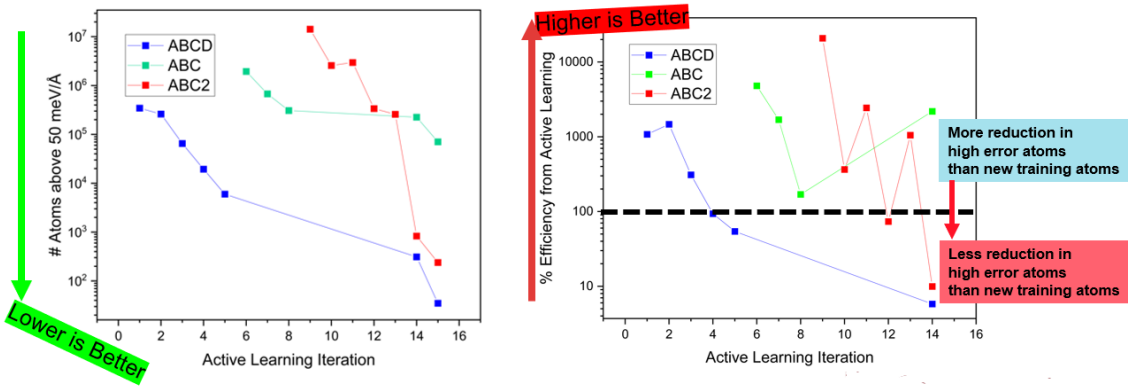

Figure 4.10 The number of atoms beyond the $50\ \mathrm{meV/Å}$ threshold and the percent efficiency from active learning.

In this work, one of the primary goals is to significantly reduce the cost facing expensive LTC calculations with pure DFT through Elemental-SDNNFF force calculation. Here, the overall cost of evaluating LTC of structures with trained NNFF models involves several steps, including model training and labeling. Figure 4.11 shows a direct comparison of the core hours needed to evaluate all 11,186 structures with the Elemental-SDNNFF and pure DFT calculations. Core hours toward Elemental-SDNNFF development include network training, labeling of training data, and labeling DFT-LTC data. The core hours for computing LTC for all structures with pure DFT is also included for comparison. The computation for the core hours in training a model is broken

down into several steps. The first cost is the labeling process of displaced structures by running DFT to obtain the atomic forces. Based on the DFT calculations presented here, each supercell approximately takes one hour on average to evaluate with 48 cores. With 620 ABCD, 1,334 ABC, and 3,822 ABC_2 supercells from active learning, this adds up to 2.77×10^5 core hours. Also, 10,298 ABCD, 11,628 ABC, and 4,435 ABC_2 supercells from DFT-LTC yields an additional 1.27×10^6 core hours. The second cost is the model training for all five models in the committee, which was approximately 20,600 core hours per training iteration. The sum across all 21 training sessions gives 4.33×10^5 core hours toward training. Adding both labeling and training together, a total of 1.98×10^6 core hours are used to develop the current Elemental-SDNNFF model. Evaluating LTC with pure DFT is estimated by 2,377 ABCD with 10 supercells each, 2,660 ABC with 18 supercells each, and 6,829 ABC_2 with 10 supercells each at one hour per supercell with 48 cores, yielding 6.72×10^6 core hours. Overall, only 29.4% of the core hours used to evaluate LTC with DFT are used in the development of the current Elemental SDNNFF model.

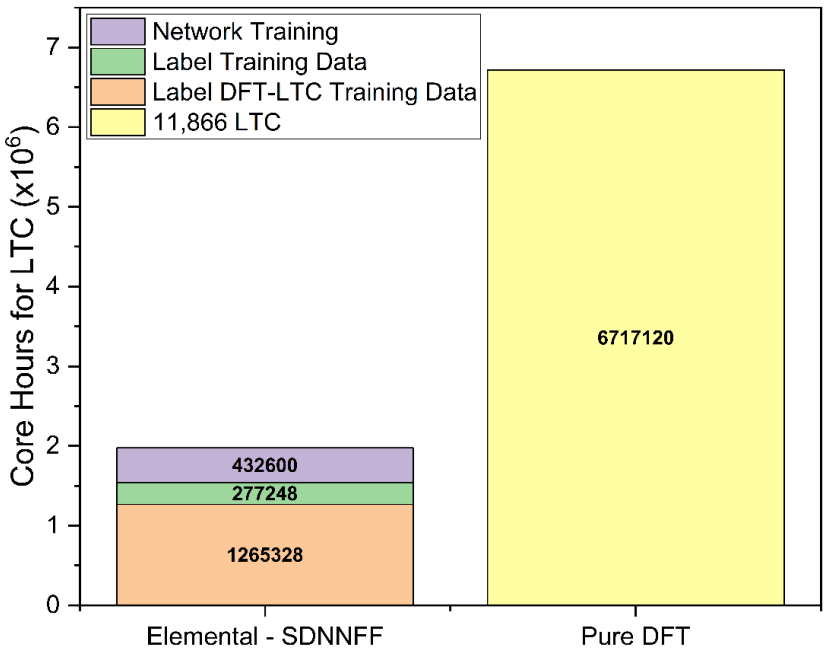

Figure 4.11 The core hours for computing lattice thermal conductivity for the 11,866 structures studied here.

4.12 Testing of the Model Force Accuracy

From the overall dataset including those from active learning and DFT-LTC, an 80%-20% training-testing split of the structures after shuffling is performed, where the remaining 20% serves as the testing set to probe the ability of the network to predict previously unseen structures. We perform the split based on *structures* and not *atomic configurations*, whereby structures and corresponding atoms in the testing set may never have been previously observed by the model. Then, as shown in Figure 4.12, four 7 Å, K=12 networks are trained on the 80% training set and a comparison between DFT and predicted

forces in the testing set is done. The first network is the original SDNNFF model labeled as the “Original SDNNFF” [96]. The jump in performance from the first to the second network is major, dropping by 62.9 meV/Å in testing RMSE. This is primarily attributed to the inability of the original SDNNFF to consider atomic species, i.e., no knowledge is trained to distinguish the 55 elements in the training set. From the second to the third network, the rotational covariance adds a very slight 0.35 meV/Å improvement to the RMSE. Finally, the added data augmentation paired with the rotational covariance bumps the RMSE to the sub-10 meV/Å mark. Overall, the 8.8 meV/Å testing RMSE with 6.7 meV/Å training RMSE shows the robustness of the network when facing new structures containing various atomic positions and species. The small difference between the training and testing RMSE also implies little overfitting in the current model.

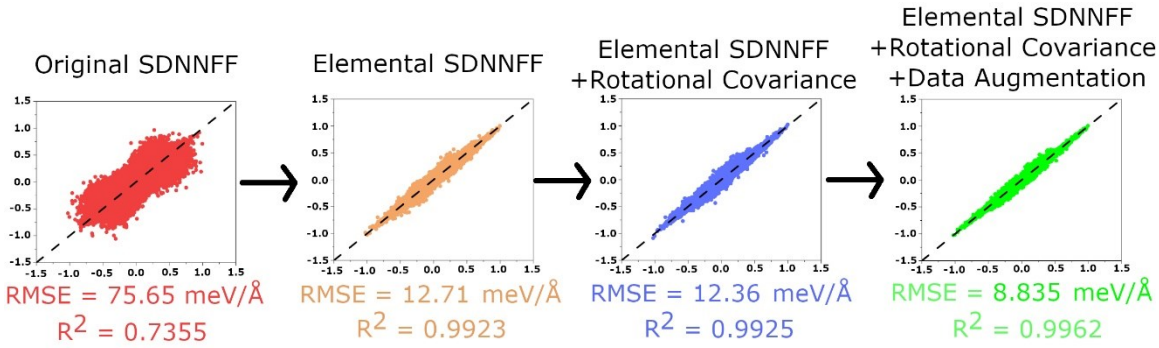

Figure 4.12 The force comparison curves between predicted and untrained values.

4.13 Phonon Frequency and Lattice Thermal Conductivity Prediction

To compare the frequencies in the phonon dispersions from Elemental-SDNNFF with those from DFT, the RMSE across all bands are computed. Since

the range of the phonon dispersions vary between structures, we divide the RMSE by the maximum phonon frequency to scale all structures equally and is expressed as a percentage. To display the agreement of phonon dispersions, two ABC_2 structures are taken from the testing pool for benchmark comparison. As seen in Figure 4.13, the overlap with DFT for $RuAuMg_2$ and $CrFeTa_2$ is virtually perfect in the acoustic branches with near-perfect quality in the optical branches. The RMSE/range percentages for $RuAuMg_2$ and $CrFeTa_2$ are 0.5459% and 0.4931%, respectively, owing the performance to the quality of the 2^{nd}-order IFCs and corresponding atomic forces. The quality of the phonon frequencies as a function of Elemental-SDNNFF cutoff are further investigated. The overall histogram containing the RMSE over maximum phonon frequency ratio for all structures is shown in Figure 4.14.a. The increase of the force cutoff improves the accuracy for the predicted frequency from the observed push of the distribution towards the left of the figure. This is expected as the harmonic force constants are sensitive to long-ranged interactions which are truncated by the finite cutoff [49], especially the existence of long-ranged dipole-dipole interactions in polar solids [142]. Additionally, the error in Table 4.1 seems to increase with decreasing average number density in the order ABCD, ABC_2, and ABC corresponding with the possible truncation of atomic neighbors beyond the cutoff.

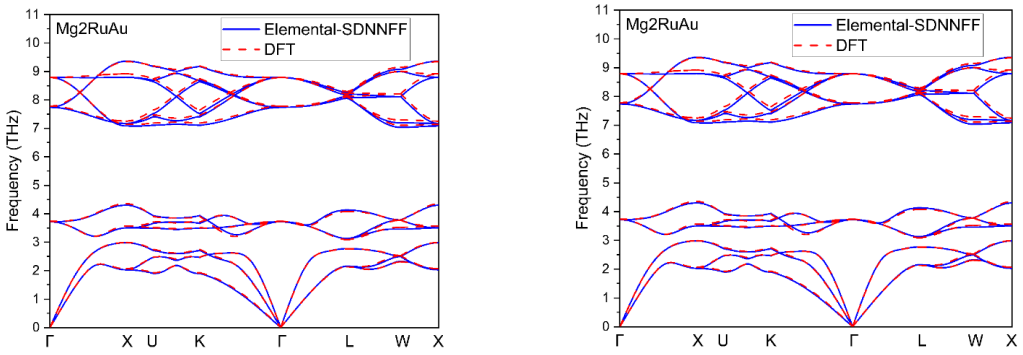

Figure 4.13 The phonon dispersion of $RuAuMg_2$ and $CrFeTa_2$ structures.

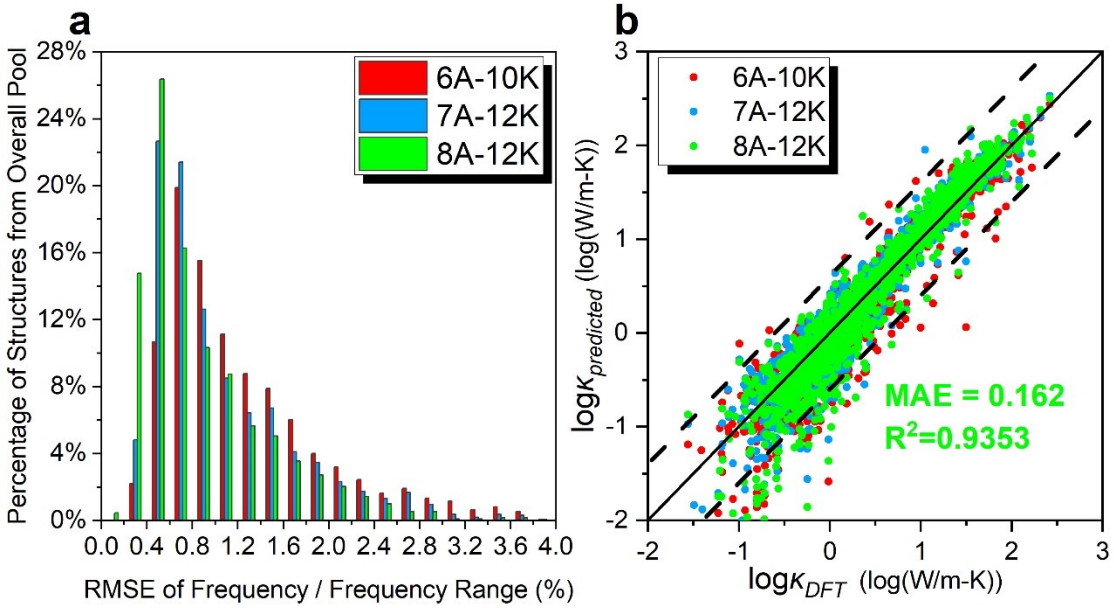

Figure 4.14 Histogram and comparison plot of predicted phonon properties on 1,298 untrained structures.

Table 4.1 The average phonon frequency error for each structure type against various networks.

Structure type	6 Å-10K	7 Å-10K	8 Å-10K
ABCD	0.94%	0.75%	0.71%

ABC	1.55%	1.11%	1.04%
ABC2	1.36%	1.13%	1.02%
Overall	1.25%	1.05%	0.90%

The LTC involves both the 2nd and 3rd-order IFCs, the latter being significantly more sensitive to the force error [127]. For verification, 553 ABCD, 649 ABC, and 583 ABC_2 totaling 1,785 structures are set aside for DFT-LTC calculation for comparison with those results from the SDNNFF. As seen in Figure 4.14.b, the agreement with DFT is very good following linear trends with observable outliers. The mean absolute error (MAE) of the LTC in the 8Å-12K (cutoff of 8 Å and k = 12) model for ABCD, ABC_2, and ABC are 0.0934, 0.123, and 0.2526 $Wm^{-1}K^{-1}$, respectively, with an overall MAE of 0.162 and R^2 of 0.9353. The former two structure types are in agreement especially when compared with the 0.12 MAE and 0.87 R^2 predicted by Zhu *et. al* [121]. The ABC structures own relatively lower accuracy, which could be improved with further active learning iterations. Further insight to the effects of the cutoff to the LTC accuracy is provided by histograms in Figures 4.15-4.17. Due to the wide range of computed LTC (~10^{-1} to 10^2 $Wm^{-1}K^{-1}$), the absolute percent difference of the LTC instead of the RMSE is used herein to quantify the performance. As seen by the histograms and Table 4.2, there is a weak correlation of absolute percent difference with respect to the cutoff. However, there is evidence of a slight leftward shift of ABC with increasing cutoff which is most likely due to the low density of ABC in which the 8 Å force cutoff is more likely to capture the

interactions up to the third nearest neighbor as mentioned previously. Overall, despite the accuracy of the 2nd-order IFCs as shown by the phonon dispersions, the 3rd-order IFCs are more challenging to capture across the many structures studied here. Nonetheless, a significant proportion of structures stay within factor of 2 of the predicted LTC as seen by both the 2× margin in Figure 4.14.b and the high population of structures within absolute percent error of 100% in the histograms of Figures 4.15-4.17. It is worth pointing out that, for ultralow LTC prediction, this could be treated as a rule-of-thumb to filter structures by, for example, <1 $Wm^{-1}K^{-1}$. Given that the prediction is within 2× DFT-LTC values, structures with predicted LTC < 0.5 $Wm^{-1}K^{-1}$ are highly expected to lie within the target 1 W m^{-1} K^{-1} range. Therefore, our Elemental-SDNNFF is promising for the filtration of ultralow LTC structures.

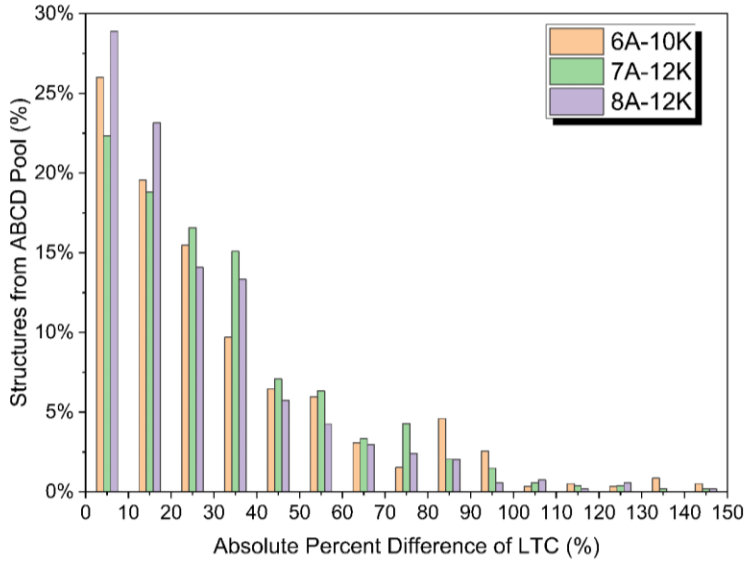

Figure 4.15 Histogram containing the absolute percent difference of lattice thermal conductivity for 553 ABCD structures.

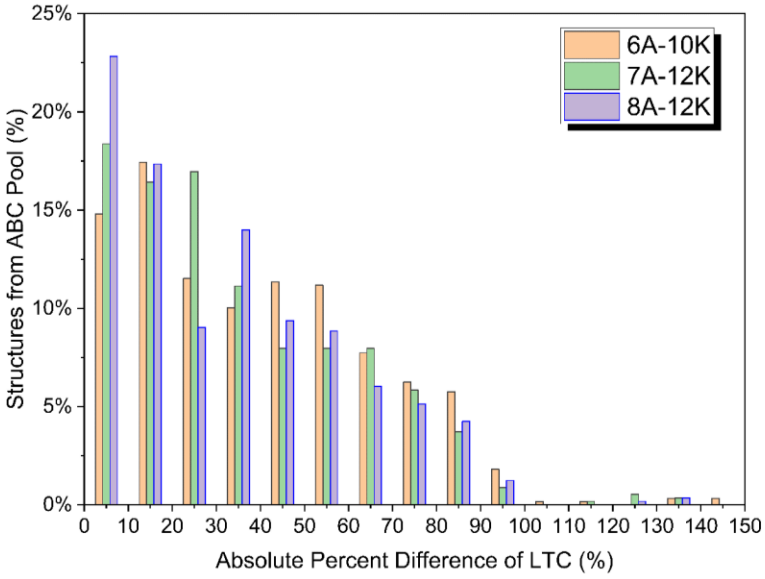

Figure 4.16 Histogram containing the absolute percent difference of lattice thermal conductivity for 649 ABC structures.

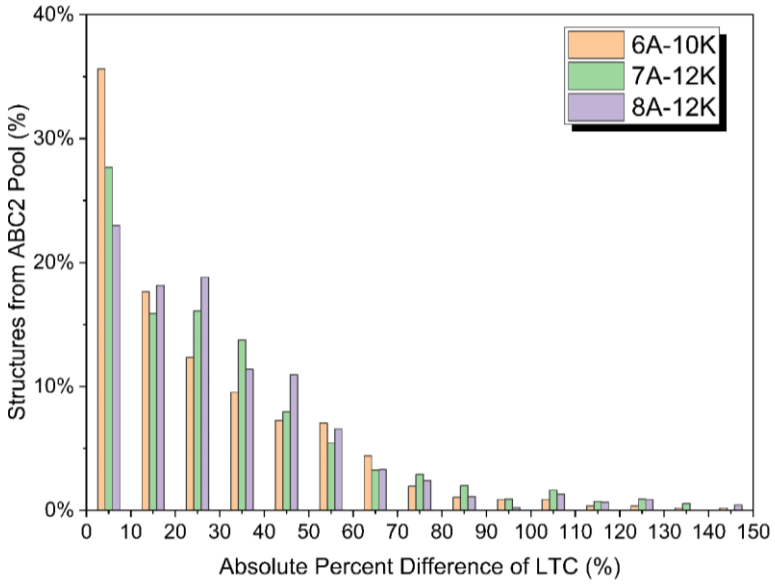

Figure 4.17 Histogram containing the absolute percent difference of lattice thermal conductivity for 583 ABC_2 structures.

Table 4.2 The average absolute percent difference for lattice thermal conductivity.

Structure type	6 Å-10K	7 Å-12K	8 Å-12K
ABCD	30.75%	32.58%	28.07%
ABC	41.20%	38.58%	39.44%
ABC2	30.59%	34.13%	31.41%
Overall	35.10%	33.90%	32.20%

Comparison of the LTC from the literature containing first-principles and experiments is provided in Table 4.3 along with our corresponding DFT and predicted values. As seen from the table, comparison of the DFT-LTC with those

from other works holds an average percent difference of 4.955% which is expected considering differences in DFT parameters such as energy convergence criterion and pseudopotential or by method of computing the force constants and associated parameters such as displacement distance and q-point mesh. Correspondingly, the agreement of our predictions with the DFT-LTC values from the literature inherits a similar 6.294% average percent difference and may be attributed to the difference in DFT parameters and corresponding forces from the training data. More importantly, the experimental values are emphasized here as the high-throughput predictive capabilities demonstrated above are designed for near-future synthesis and deployment of crucial thermal materials. Due to the little abundance of experimental data and large coverage of the structures studied here, only one structure with experimental LTC found in the literature was not once trained into our model. The structure is ZrNiSn whereby the average experimental LTC is 11.5 $Wm^{-1}K^{-1}$ [143] with a literature DFT value of 10 $Wm^{-1}K^{-1}$ [144] whereas this work's DFT and prediction values are 15.1 and 14.5 $Wm^{-1}K^{-1}$, respectively. The existing difference in the experimental LTC with respect to DFT and prediction values is attributed to effects not considered in solving phonon BTE, such as defects, boundary scattering, phase separation, and electron-phonon interactions [145–147]. Because these effects are collective in defining the phonon scattering rates and therefore detract from the overall LTC, almost all the predictions *via* DFT or Elemental-SDNNFF are expectedly slightly higher than experiment. Nonetheless, the filtration of high-performance thermal materials given this knowledge remains

highly feasible provided the error is akin to that between DFT and the predictions for the entire pool of Heusler structures studied here. Indeed, the model presented here captures the forces and corresponding LTC for the previously unseen material ZrNiSn displaying successful prediction for unseen structures with unique composition and lattice size.

Table 4.3 List of structures and their lattice thermal conductivities compared with previous studies.

OQMD id	Compound	DFT (This work)	Prediction	DFT (Other works)	Experiment
947107	AuAlHf	26.511	17.914	16.7 [128]	-
860047	CoAsHf	16.584	22.904	20 [128]	-
977758	BiKSr	2.0764	0.97076	1.96 [128]	-
859488	ZrCoAs	18.479	24.073	24 [128]	-
861248	ZrCoSb	20.999	26.43	25 [128]	15 [148]
859303	TaCoSi	27.587	35.531	37.8 [128]	-
959187	TiNiPb	19.47	22.211	109 [128]	-
960180	ScNiSb	-	11.975	19.5 [128]	-
573181	TaSbOs	22.4	22.572	29.6 [128]	-
900042	YAsPd	-	2.6595	5.48[128]	-
566964	TaGaPt	27.4	27.115	32.9[128]	-
567451	TiGePt	20.7	24.766	16.9 [128]	-
885549	HfSbRh	-	15.691	21.8 [128]	-
933731	VAsRu	23.2	20.662	23.5 [128]	-
917433	KCaSb	-	0.2577	2.7 [128]	-

580774	SrCdSi	-	0.71576	13.5 [128]	-
924394	BaSrSn	-	0.91322	2.01 [128]	-
845211	TeAgLi	1.62	0.56017	1.52 [128]	-
977865	KBaBi	-	0.17887	2.19 [128]	-
978542	SrLiBi	-	0.39808	3.04 [128]	-
859488	ZrCoAs	19.8	24.073	24 [128]	-
860757	HfCoBi	-	15.409	18.6 [128]	-
860547	TiCoSb	-	21.475	-	18 [149], 25 [148]
958939	HfNiSn	-	14.218	8.5 [144]	8.5 [150]
861248	ZrCoSb	21.1	26.43	-	20 [146]
958269	TaFeSb	25.428	28.634	-	8.8 [151]
959183*	ZrNiSn	15.079	14.468	10 [144]	11.5 [143]
115027	NiTiSn	13.884	14.35	15.5 [145]	7.5 [152]
647222	VFeSb	20.99	24.116	18.6 [153]	12.2 [153]
18705	AlVFe2	61.373	66.855	-	22 [154]
958416	NbFeSb	23.956	34.548	-	16 [155], 14 [151]

*Untrained structure.

4.14 t-SNE Analysis

Currently, the atomic weight of central atom, the neighboring atomic positions, and neighboring species serve as the input to the Elemental-SDNNFF yielding the atomic forces on the central atom. These atomic forces are not directly related to the phonon properties but are necessary to construct the IFCs and

subsequently solve phonon BTE. As such, the relationship between atomic-level information and global properties such as the LTC is not straightforward. Here, we instead pursue a higher-level understanding of phonon properties *via* t-distributed stochastic neighbor embedding (t-sne) method. The t-sne is a dimensionality reduction method allowing visualization of highly complex vectors into 2D/3D points, whereby the proximity of these points defines their correlation. To understand the Elemental-SDNNFF input vector, reduction to 20 dimensions by principal component analysis (PCA) [156] followed by further reduction to 2D by t-sne is performed on the entire pool of structures with predicted positive dispersions (7,373). Additionally, because each structure contains several atoms, the SDNNFF mesh is instead centered about the entire primitive cell as opposed to each atom. This guarantees only one vector input per structure is plotted for t-sne analysis. Then, the points are colored based on global properties, including structure type, cell volume, number density, average atomic mass, mass density, and the predicted LTC to observe any structure-property relationships in Figure 4.18. As shown in the figure, colors are added for every six plots based on structure type (red is ABC_2, green is ABC, and blue is ABCD), volume ($Å^3$), number density (atoms per $Å^3$), average atomic weight (amu), mass density (g/cm^3), and predicted LTC ($\log(Wm^{-1}K^{-1})$).

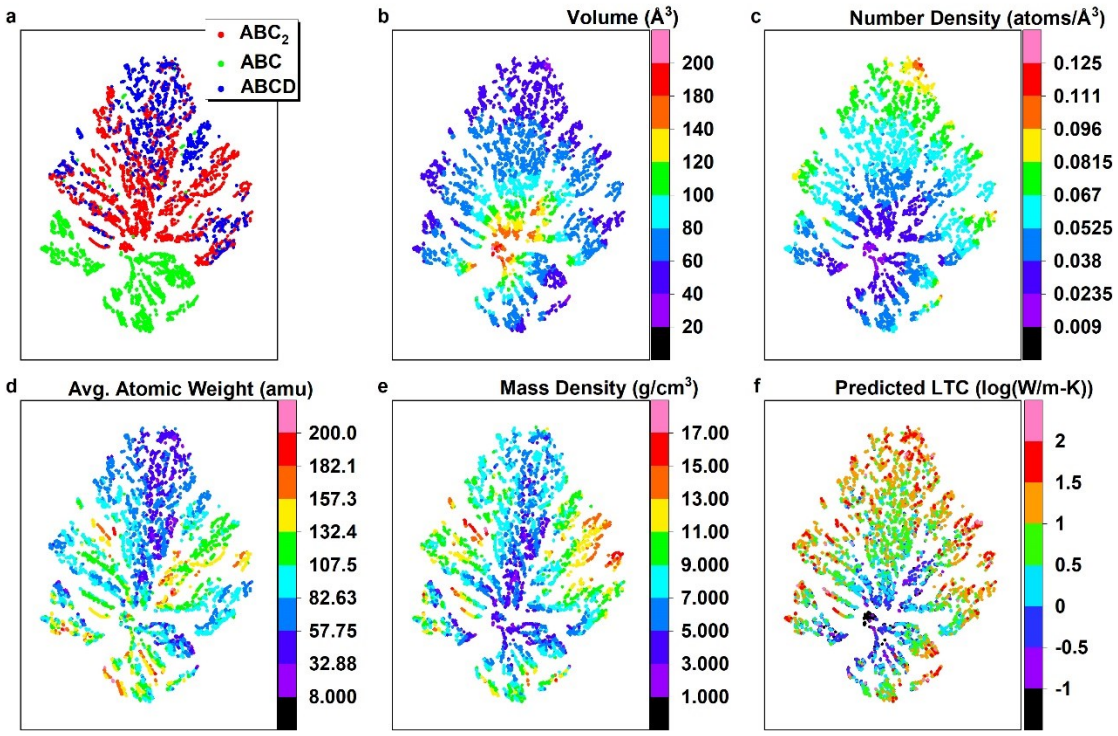

Figure 4.18 t-sne plots of model inputs colored by the respective property.

As seen in Figure 4.18.a colored based on structure type, the green ABC points are distinguishable from the mixed red ABC_2 and blue ABCD points. Indeed, ABC structures, when compared with ABC_2 and ABCD, own a missing lattice site, causing isolation of ABC structures as they are clearly identified by the Elemental-SDNNFF input. On the other hand, the ABC_2 and ABCD structures here share similar lattice sites with only difference lying in different elements on the lattice sites (ABCD structures have additional element than corresponding ABC_2 structures and thus have lower symmetry), and thus some overlap is expected. Interestingly, Figures 4.18.b and 4.18.c clearly show a plateau for the primitive cell volume and a valley for the number density at the center of the t-sne

bordering the ABC and ABC_2/ABCD clusters. Despite differences in the structure, all structure types are involved with the same gradients, i.e., a decrease in volume and increase in number density approaching the outer edges of the plot. Indeed, the relative positions of atoms are expected to expand or contract depending on their occupied elements and corresponding bonding behavior. Thus, because the Elemental-SDNNFF captures the atomic positions with a constant cutoff, the volume and number density are characterized per structure. Figures 4.18.d and 4.18.e are colored based on the average atomic weight and mass densities sharing similar distributions at a quick glance. Because these values involve summing over all atoms in the primitive cell, it is not expected that the t-sne plots own obvious trends like the volume or number density since the Elemental-SDNNFF considers the contribution of each atom individually. Nonetheless, these two plots offer some insight to the distribution of the dataset studied here. For instance, many of the ABC_2/ABCD structures own either extremely low or high average atomic masses and mass densities, whereas ABC is somewhere in-between. Also, many of the structures with low volumes at the center also correspond with low mass density which is sensible.

The most interesting result is seen in Figure 4.18.f whereby the color is based on the LTC predicted by our Elemental-SDNNFF model. As seen from the plot, the ultralow LTC structures (<1 $Wm^{-1}K^{-1}$, corresponding to zero in logarithm scale) are toward the center of the plot and the high LTC structures exist toward the outer edge. This marks an interesting direct relationship with the number density and inverse relationship with the total volume, which is consistent with

Zhu *et. al.* in which volume-related features like average atomic ground state volume, average bond length, and volume per atom own high feature importance in the LTC [121]. In general, ultralow LTC structures are significantly more distinguishable when compared to the high LTC structures which are more scattered. Nonetheless, general conclusions could be made to categorize the thermal transport performance of these structures based simply on the number density. For instance, when comparing the number density and LTC plot, structures with 1 $Wm^{-1}K^{-1}$ or less are expected to own a number density up to ~0.0525 atoms/$Å^3$, whereas structures with more than 30 $Wm^{-1}K^{-1}$ are expected to own a number density of at least 0.067 atoms/$Å^3$. Note, these estimates should only apply for the ABCD, ABC_2, and ABC Heusler structures studied herein. Additionally, a Pearson correlation plot is shown in Figure 4.19 showcasing the strong positive correlation (+0.71, red) and negative (-0.64, blue) correlation of the LTC with respect to number density and volume of primitive cell, respectively, whereas the other two properties show weak correlation as summarized by the t-sne plots. As validation, Figure 4.20 provides a colored comparison between the predicted LTC and DFT-LTC for the 1,298 structures serving as the test set. This is the same t-sne plot as before but with all other structures than those in the test set excluded. As seen by Figure 4.20, the trends between the predicted and DFT-LTC are matching with ultra-low LTC in the middle and generally showing an increase in LTC as one extends to the outer surface formed by the t-sne points. This shows that the predictions made by the Elemental-SDNNFF can represent DFT trends and shows the potential for high-

throughput evaluation of materials databases. Overall, simple descriptors such as number density are extremely valuable for quick filtration of structures in databases as they require little to no cost to compute. Discovery of trends with respect to LTC are made possible by the rapid prediction capabilities of the Elemental-SDNNFF presented here, which we expect to expand to other structure types in the near future.

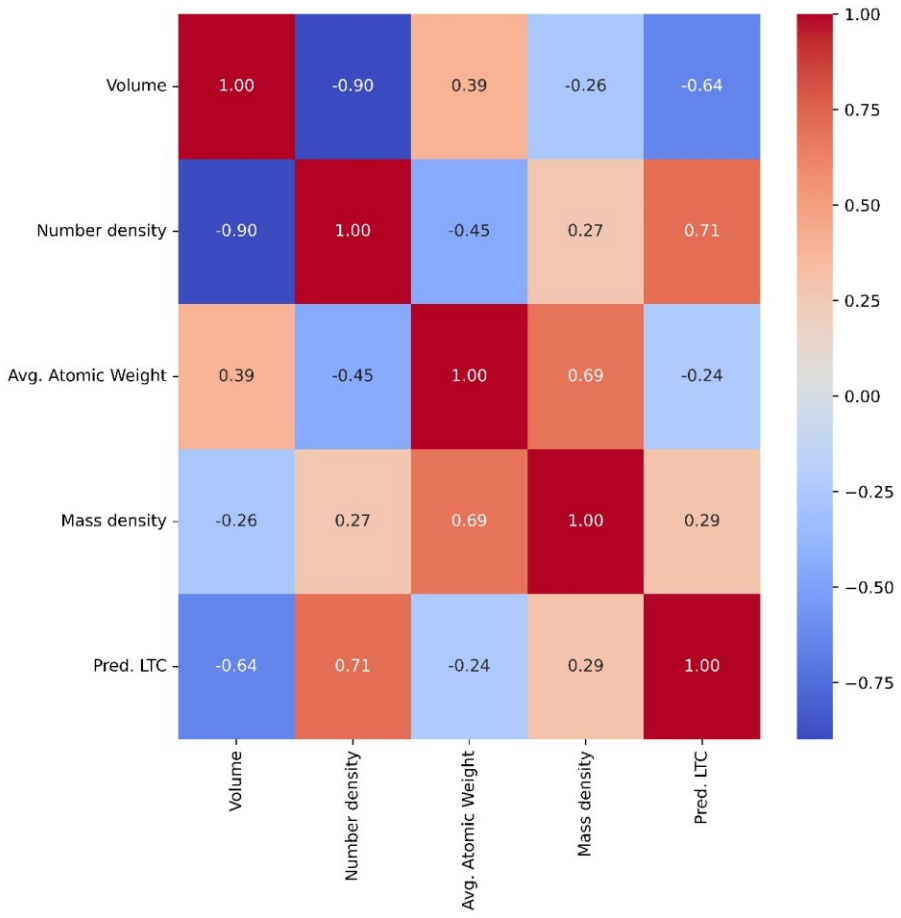

Figure 4.19 Pearson correlation plot of properties studied here.

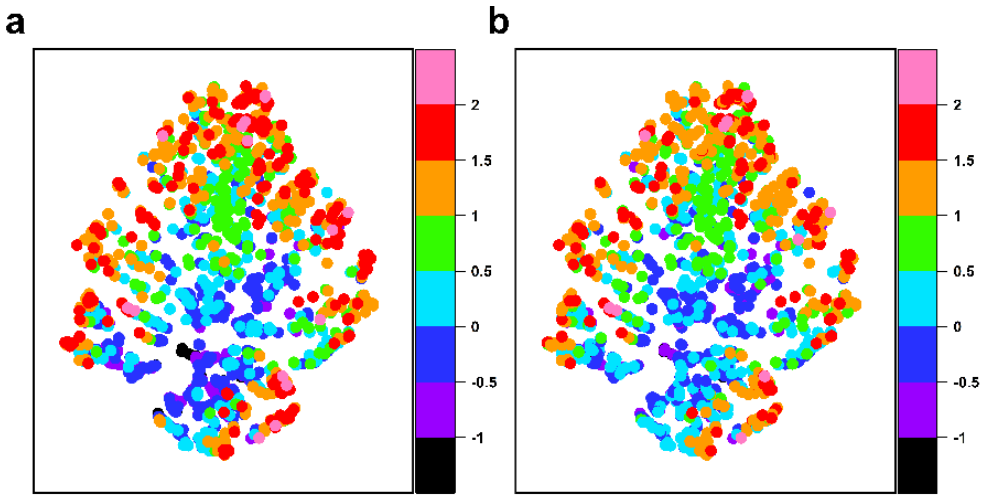

Figure 4.20 t-sne plots of the testing set.

4.15 Insights into Phonon Anharmonicity from p-d Orbital Hybridization

With the LTC of 7,373 out of 11,866 thermodynamically stable Heusler structures with additional 1,298 structures from the testing set accurately predicted, we are now in a position to provide deep insight into the phonon anharmonicity of crystalline materials. The mean atomic mass has long been regarded as a good predictor for LTC of several materials as Slack and Keyes formula state [157,158]. The higher the mean atomic mass with the same number of atoms between two materials, the lower LTC is and vice versa [157,158]. However, there are some exceptions that materials having same atomic mass might have quite different LTC due to various anharmonicity mechanisms other than mean atomic mass. For example, CuBr, ZnSe, and GaAs have the same number of atoms (2 atoms per primitive cell, denoted as M-X with M for cations

and X for anions) and mean atomic mass of ~ 72 amu. However, CuBr, ZnSe, and GaAs possess different LTC of 1.25, 19, and 45 $Wm^{-1}K^{-1}$ at 300 K, respectively [159]. Such phonon anharmonicity mechanism was discovered by Jaffe et al. [160] and Wei et al. [161], which states that the hybridization in M-d orbitals and X-p orbitals can cause repulsion. The strength of the hybridization is reflected by the overlap and energy difference between the orbitals in M-d and X-p [159–161]. Moreover, the overlap and hybridization between M-d and X-p causes antibonding states below Fermi level which causes more anharmonicity in the material [159]. We examine the phonon anharmonicity caused by p-d hybridization for the trained full-Heusler structures. Two materials, namely Li_2PdAs and Li_2CdGa, were selected from the Elemental-SDNNFF model as candidates for in-depth analysis with explicit DFT. One particular reason why these two materials are selected is because they have the same number of atoms in the primitive cell and almost same mean atomic mass of ~49 amu. However, they have different values of LTC (1.92 $Wm^{-1}K^{-1}$ for Li_2PdAs and 3 $Wm^{-1}K^{-1}$ for Li_2CdGa) which ensures that the difference in LTC is caused by an anharmonicity mechanism other than the mean atomic mass. In fact, the p-d hybridization phenomenon is observed in Li_2PdAs with reasonable p-d hybridization, while Li_2CdGa has weak or non-existent p-d hybridization in Figure 4.21. Figure 4.21.a shows the projected band structure for d orbitals in Pd and p orbitals in As in Li_2PdAs, while Figure 4.21.b provides the Crystal Orbital Hamilton Population (COHP) for Pd–As bond in Li_2PdAs (negative COHP represents antibonding and positive COHP represents bonding). In Figures

4.21.c and 4.21.d, the projected band structure for d orbitals in Cd and p orbitals in Ga and the COHP for Cd-Ga bond in Li_2CdGa are given, respectively. The Fermi energy is scaled to 0 eV in all the subplots. Since Li in both materials has neither p-orbitals nor d-orbitals, the analysis on orbitals is then only done on d-orbitals in Pd/Cd and p-orbitals in As/Ga in the orbital-projected band structures. The Crystal Orbital Hamilton Population (COHP) was also calculated in both materials to further explain the bonding and antibonding states especially the state below Fermi level between Pd/Cd and As/Ga.

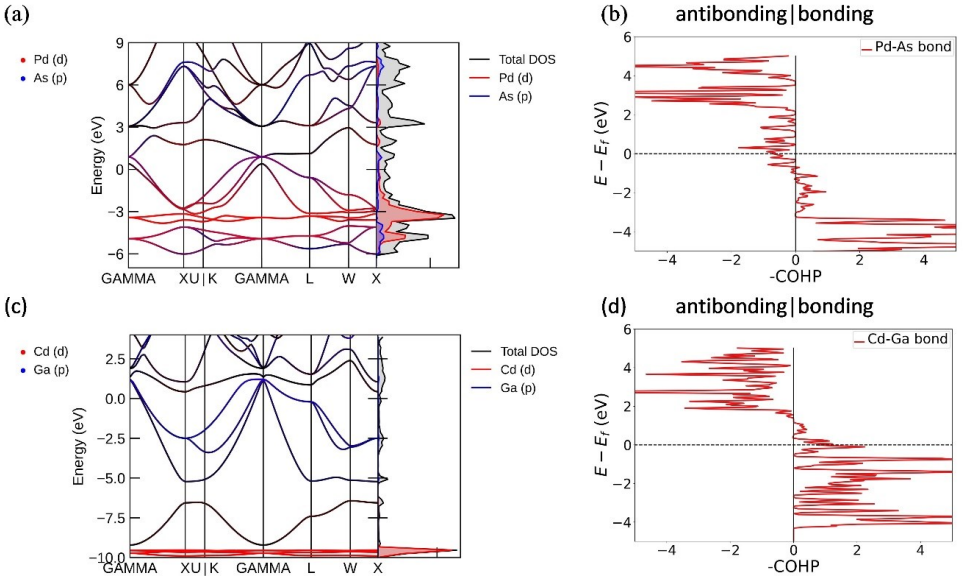

Figure 4.21 Electronic band structure and COHP analysis of Li_2PdAs and Li_2CdGa.

The difference between LTC can be explained by the Pd/Cd-d orbitals and As/Ga-p orbitals hybridization [160,161]. The orbital-projected band structure of Li_2PdAs shown in Figure 4.21.a confirms the presence of hybridization between d-orbitals in Pd and p-orbitals in As. The colors red (d-orbitals in Pd) and blue (p-

orbitals in As) below Fermi level overlap with each other in Figure 4.21.a which indicates the hybridization between red (d-orbitals in Pd) and blue (p-orbitals in As). Moreover, the electronic density of states in Figure 4.21.a shows the overlap between the red d-orbitals in Pd with the blue p-orbitals in As. In Figure 4.21.b, the presence of antibonding negative COHP between Pd-As bond in Li_2PdAs below Fermi level also confirms the presence of p-d hybridization in Li_2PdAs between Pd-d orbitals and As-p orbitals. In Figure 4.21.c the orbital-projected band structure of Li_2CdGa, the red (Cd-d orbitals) and blue (Ga-p orbitals) colors do not overlap in the orbital-projected band structure. Also, the density of states shows no overlap between red Cd-d orbitals and blue Ga-p orbitals partial density of states which confirms the weak or non-existent hybridization between Cd-d orbitals and Ga-p orbitals. In Figure 4.21.d, the non-existent antibonding (negative COHP) states below Fermi level in Cd-As bond confirms the lower anharmonicity in Li_2CdGa than in Li_2PdAs. The difference in LTC may not be significant, which is due to the fact the p-d hybridization is weaker in Li_2PdAs compared to the p-d orbitals hybridization in CuBr that occur between Cu-d orbitals and Br-p orbitals [159]. Despite the small hybridization between Pd-d orbitals and As-p orbitals in Li_2PdAs, it is still worth taking it into consideration since it unravels a new understanding on phonons anharmonicity in materials caused by orbitals, which is expected to help design materials with ultralow LTC for vast applications such as thermoelectrics and thermal insulation.

To further analyze the chemistry effect on the LTC, Figure 4.22 displays the bonding vs. antibonding for all 11,866 Heusler structures studied herein.

Specifically, Figure 4.22.a contains the testing set structures colored by their corresponding DFT-LTC, whereby Figure 4.22.b includes all other structures colored by their LTC predicted by our Elemental-SDNNFF model. The bonding and antibonding value for each crystal structure is obtained *via* integration of the COHP curve. When both figures are compared, the general trends are matching between DFT and predicted LTC and supports their agreement. One notable trend is seen by the group of structures at low bonding (< 200) and high antibonding (> 1). Here, only low LTC structures $<10^{0.5}$ $Wm^{-1}K^{-1}$ are observed and agrees with the earlier discussion of highly anharmonic materials containing antibonding behavior. As seen by the insets, the logarithm of bonding value is inversely related to the volume, and as the bonding decreases or volume increases, both DFT and predicted LTCs also generally decrease. This corresponds with the earlier finding in Figure 4.18 whereby the volume is inversely proportional to the LTC. In essence, the high antibonding and low bonding behavior may be proposed as a method for quick filtration of insulating crystals necessitating only the unit cell in the DFT calculation. On the other hand, another trend is seen at the high bonding region (> 200). Observably, the entirety of the LTC range is found here despite the antibonding, indicating the competing bonding-antibonding behavior for the LTC. However, only beyond this bonding region will high LTC $>10^{0.5}$ $Wm^{-1}K^{-1}$ ever be observed, which may prove useful for filtration of high performance thermally conductive crystals.

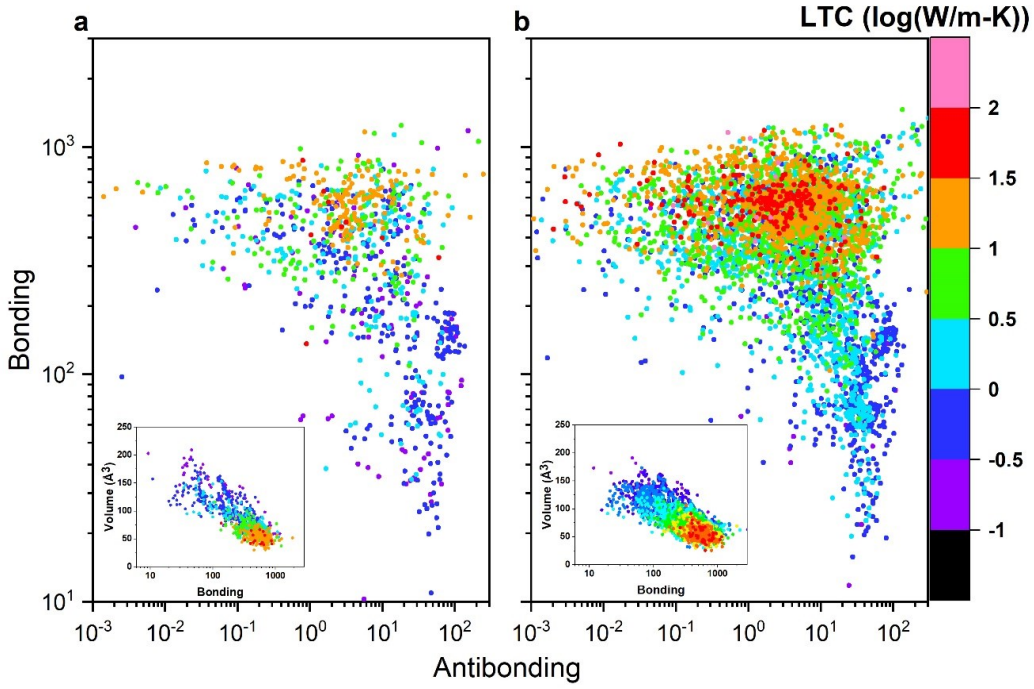

Figure 4.22 Bonding vs. antibonding colored by lattice thermal conductivity.

4.16 Weyl Points Prediction

Until now, we have focused on phonon properties related to their eigenvalues, while topological effects of phonons originate from the correlation of their eigenvectors. Nevertheless, the topological classifications of phonons are obtained from harmonic force constants. Here, harmonic force constants from the Elemental-SDNNFF are used to perform the search due to their relatively high accuracy to DFT. Figure 4.23.a contains the ratios of topological Weyl semimetals in ABC and ABCD structures. An extensive searching for WPs is performed on space group number 216 Heusler structures, revealing that 68.7% of ABC structures and 87.6% of ABCD structures have WPs as seen in Figure 4.23.a. Such high successful rate is much higher than the recent high-throughput

study of electronic materials which found only 30% are topological [162]. Since all frequency range of phonons can be stimulated, this further demonstrates the advantage of phonons as a platform for studying topological states. Figure 4.23.b provides the number of structures containing different types of WPs, including WPs with high symmetry (on XW, on ΓXWK, and on XWU), WPs without high symmetry (in bulk), clean WPs, and double WPs. All WPs are categorized according to their symmetry, including those on high symmetry lines, on high symmetry surfaces and in bulk. The insets illustrate clean WPs and double WPs. Figure 4.23.c shows the plot of first Brillouin zone and irreducible Brillouin zone of space group 216. Although the concept of metal and insulator breaks down when describing phonons, a clean semimetal in which the density of states vanish at gap closing point still benefits experimental verification. We found 85 and 92 clean WPs in ABC and ABCD Heusler structures, respectively, while previous study has demonstrated clean WPs in half-Heusler materials [163]. Clean WPs are found between acoustic and optical branches, and between clusters of optical branches. We have found clean WPs between band 3/4 and band 6/7 in ABC structures and between band 3/4, band 6/7 and band 9/10 in ABCD structures. Interestingly, another class of two-band charge-2 WPs referred to as ‘double WPs’ is encountered, which has not been discovered in phonon systems previously. The double WPs are accompanied by a pair of chiral Fermi arcs, expecting novel transport properties. They also serve as parents of other topological phases, giving birth to two charge-1 WPs upon symmetry breaking. The double WPs predicted in this work may offer guidance on topological phase

transition experiments. Note that there are also three-band charge-2 WPs and four-band charge-2 WPs [131], but they are not included in this work because of the symmetry constrains of the Heusler structures. Unlike linear WPs with Chern number ±1, the double WPs have Chern number ±2 and exhibit quadratic dispersion near the band-crossing point [164]. The double WPs are listed in Tables 4.4-4.5. The positions of all WPs are shown in Figure 4.23.d-4.23.g, showing the positions of WPs in bulk, on XWU, on ΓXWK, and on XW, respectively. Since our dataset includes 55 elements spanning the periodic table, our result can be viewed as a full evaluation of all possible WPs in Heusler structures. Most of WPs appear on high symmetry line XW and mirror plane ΓXWK, and the distribution is extensive. The frequencies of WPs in quaternary ABCD Heusler structures are found to be higher than those in ternary ABC Heusler structures, due to more atoms and denser packing in the unit cells of quaternary structures.

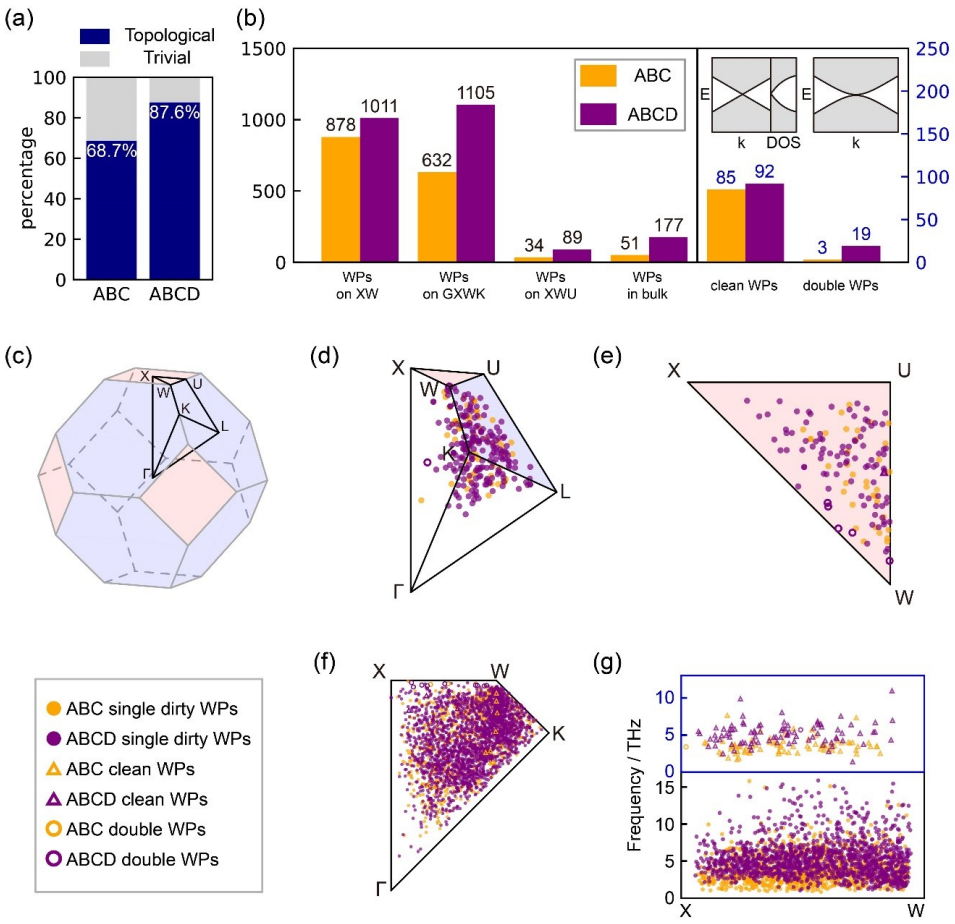

Figure 4.23 Results of searching Weyl points in 3,212 Heusler structures with phonon dispersions predicted by the model.

Table 4.4 List of space group 216 ABC structures discovered with double WPs.

OQMD id	formula	k1	k2	k3	lower band	frequency	chirality
869233	TlInGe	0.492	0.136	0.628	2	1.1166	2
861291	NbCoSb	0.49	0.094	0.584	1	2.2066	2
932904	ZrBeAl	0.5	0.005	0.505	2	3.3941	2

Table 4.5 List of space group 216 ABCD structures discovered from OQMD with double WPs.

id	formula	k1	k2	k3	lower band	frequency	chirality
794976	YInCuAg	0.487	0.221	0.708	5	2.9014	2
712368	ScZnGaCu	0.509	0.203	0.695	7	3.9725	2
1025695	LiScBiAu	0.512	0.173	0.661	2	1.2489	2
789193	ZrInCuPd	0.484	0.053	0.537	4	3.2282	-2
800197	NbAlFeRh	0.514	0.249	0.735	2	3.4666	2
817894	ZrZnNiRu	0.272	0.19	0.462	5	4.7757	-2
765032	NbAlCoRu	0.503	0.186	0.683	6	5.4609	2
804614	AlFeMoOs	0.488	0.227	0.715	5	4.68	2
998423	LiTiSnRh	0.489	0.2	0.689	6	5.7335	-2
810682	HfMnSnRh	0.495	0.048	0.543	3	3.3142	2
765181	HfInCoRu	0.487	0.086	0.573	2	3.3189	2
996860	LiVGaRh	0.492	0.128	0.62	6	6.1472	-2
798818	LiZrFeSn	0.314	0.084	0.372	4	4.5358	-2
997120	LiHfSnRh	0.488	0.205	0.693	3	3.2736	2
810610	TiMnAlRh	0.5	0.123	0.623	4	5.6939	2
781875	MnCrSiMo	0.489	0.073	0.562	7	6.5374	-2
702041	LiFeSiNi	0.395	0.353	0.73	4	6.1214	-2
803210	MnAlFeMo	0.51	0.173	0.663	6	6.5165	2
697382	TiAlFeCo	0.506	0.223	0.703	6	6.9059	-2

4.17 Conclusion

In summary, we developed and trained a single deep neural network model dubbed Elemental-SDNNFF for predicting complete phonon properties of crystals with demonstrated high force accuracy and speed facing both observed and new compositions of half, quaternary, and full Heusler structures. Benefited from the modified algorithm that enables million-scale atomic environments as training data, the accuracy of the predicted full phonon properties (phonon dispersions and LTC as case studies) reflect the force accuracy with respect to DFT mimicking realistic electronic-level surfaces generated by atomic vibrations. The primary interest of the Elemental-SDNNFF is the capability of predicting full phonon properties of crystals in a single deep neural network model and the sustained DFT-level forces facing large-scale structures with substantial combinations of elemental compositions, effectively capturing bonding behaviors spanning the periodic table. Specifically, this is attributed to training directly on forces procuring an $N \times t$ dataset, where t is the number of DFT simulations and N is the number of atomic force vectors. Data augmentation further elevates the available data by 3× for the Heuslers studied herein providing the neural network extremely abundant data for generalizing *ab initio* force fields. Additionally, we incorporated active learning allowing the models to prioritize those previously unseen structures with the highest error for subsequent training and is shown to drive improvement understanding atomic environments undergoing lattice vibration. Made possible by the rapid evaluation of complete phonon properties of 11,866 materials, we realized the behaviors of LTC trends based on physical

and chemical properties. Mainly, we find a direct correlation of the number density with the LTC, whereas a more complex relationship with bonding and antibonding is pinpointed. For instance, high antibonding and low bonding houses ultralow LTC structures with p-d orbital hybridization, whereby antibonding is observed below the fermi level generating high anharmonicity and low LTC. Moreover, given the expansive set of force constants evaluated by the model, novel physics are also discussed in-light of Weyl points yielding several new structures containing double Weyl points which provide unique topological phonon properties. Ultimately, our work is a medium for high-throughput evaluation and quantification of full phonon properties of large-scale materials to discover phononic crystals with exceptional or tailored phonon properties unravelling insightful physics for broad materials research.

Acknowledgements

A.R. acknowledges the financial support by the Department of Energy, Office of Nuclear Energy, Integrated University Program Graduate Fellowship (IUP) under award no. DE-NE-0000095 and NASA SC Space Grant Consortium REAP Program. H.Y. and B.C. acknowledges the financial support by National Natural Science Foundation of China (51825601 and U20A20301). Research reported in this work was supported in part by NSF under awards 1905775, 2030128, and 2110033.

Author Contributions

M.H. conveyed the idea and designed and supervised the study. A.R. performed the neural network potential training and testing as well as the active learning loop. C.L. wrote the code utilizing compressive sensing lattice dynamics (CSLD) for 2nd and 3rd IFCs fitting. H.Y. wrote the code for searching Weyl points in Heusler structures. M.A. performed the COHP calculations and analysis to phonon anharmonicity. A.R., C.S., and K.C. performed DFT calculations. Y.Z. and J.H. prepared and analyzed the original structures. A.R. prepared the draft of the manuscript. K.C., J.H., B.C., H.Z., and M.H. revised the manuscript. All the authors contributed to discussions and interpretation of results in the manuscript.

Competing Interests

The authors declare no competing interests.

Chapter 5 Unlock Phonon Properties of 80,000 Cubic Crystals by an Indirect Bottom-up Machine Learning Approach[4]

[4] Rodriguez, A., Lin, C., Shen, C., Yuan, K., Al-Fahdi, M., Zhang, X., Hongbin, Z., and Hu, M., “Unlock Phonon Properties of 80,000 Cubic Crystals by an Indirect Bottom-up Machine Learning Approach,” Submitted to npj Computational Materials, 02/19/2023.

Abstract

Although first principles based anharmonic lattice dynamics is one of the most featured methods to obtain phonon properties, such method is impractical for high-throughput search of target thermal materials. We develop an elemental spatial density neural network force field as a *"bottom-up"* approach to accurately predict atomic forces of ~80,000 cubic crystals spanning 63 elements. The primary advantage of our indirect machine learning model is the accessibility of phonon transport physics at the same level as first principles calculations, allowing simultaneous prediction of comprehensive phonon properties from a single model. Training on 3,182 first principles data and screening 77,091 unexplored structures, we identify 13,461 thermodynamically stable cubic structures with ultralow lattice thermal conductivity below 1 $Wm^{-1}K^{-1}$, among which 36 structures are validated by DFT calculations. By analyzing the abundant predicted data, we propose mean square displacements and bonding/antibonding as two low-cost descriptors to ease the demand of expensive first principles calculations for fast screening ultralow thermal conductivity. Our model also quantitatively reveals the correlation between off-diagonal coherence and diagonal populations and identifies the distinct crossover from particle-like to wave-like heat conduction in various structures. Our algorithm is promising for accelerating discovery of novel phononic crystals for emerging applications, such as heat dissipation in electronics, thermoelectrics, superconductivity, and topological phonons for quantum information technology.

5.1 Introduction

The dynamics of phonons, the quanta of lattice vibrations, play a critical role in various technologies ranging from heat dissipation in modern semiconductors [8] to thermal barrier coatings in turbine blades [165]. In general, applications involving heat transfer require either extreme phonon impedance or conduction, which is often the performance-limiting property. Another application example of phonons is quantum communication, which can be realized by microwave-frequency phonons such as acoustic resonators [166] and is of significant interest for the generation of remote entanglement and the secure transmission of information. Over the recent years, material scientists have sought new materials with excellent phonon properties through a combination of laboratory synthesis and computational prediction using density functional theory (DFT). Although the former requires decades of trial-and-error with the intuition of experienced chemists [31], the latter has progressed the discovery of new materials and the understanding of microscopic phonon transport due to the availability of high-performance computers. Indeed, tremendous amount of previous studies in the thermal science have been dedicated to bulk, interfacial, low-dimensional, and layered materials [47,167–171]. Generally speaking, many investigations on thermal transport in materials have uncovered the structure-thermal property relationship, including those from bond strength, structure, and chemistry [117,172–177]. Additionally, unique phonon interactions may arise from certain crystals promoting anharmonicity and lowering lattice thermal conductivity (κ_L), including rattler atoms [1,177,178], ferroelectric instability [2], and electron-

phonon interactions [179,180]. Notably, some studies also advance the theoretical description of phonon transport, e.g. addition of fourth order anharmonicity [62] and temperature-dependent effective potential (TDEP) for temperature effects on the interatomic force constants (IFCs) [124,173,181].

Despite the robust nature of DFT in predicting thermal transport properties, the explicit treatment of electronic degrees of freedom entails significant computational costs when faced with potentially thousands of candidates to screen target phonon properties. By majority, previous studies are limited to one or few materials that may acquire results in a matter of a few days, depending on their hardware resources. However, with the advent of materials genome such as Materials Project [31] and Open Quantum Materials Database (OQMD) [182], the DFT evaluation of phase stability and properties of interest for thousands to millions of previously unexplored materials puts several years of delay on the synthesis of novel materials. This is especially magnified for phonon properties. For instance, the κ_L, one of the most important phonon properties, is computationally demanding by DFT, due to the required calculations of large amount of supercells with different atomic displacements which is then processed to give IFCs for the Boltzmann Transport Equation (BTE) simulation [49].

In response, data-driven techniques such as machine learning (ML) have surfaced in materials science to address the demanding costs of DFT, effectively trading some accuracy for significant speed-up. The basic assumption with ML-based models for predicting DFT-level properties is the introduction of a finite cutoff, whereby atomic interactions beyond such cutoff are neglected. This allows

for linear scaling of the computational cost against the number of atoms as opposed to the cubic scaling with DFT. With ML, prediction of target properties requires a physically informative set of inputs as the so-called “descriptors”. For a sufficiently accurate model, descriptors should satisfy several requirements, including a) distinguishable representation for each system, b) descriptive of the similarities/discrepancies between systems, c) completeness to sufficiently differentiate systems, and d) simplicity of the descriptor to ease calculation time [183]. Given these requirements, several methods have been proposed over the past few years to confront computationally costly phonon properties. For example, progresses have been made in the ML prediction of κ_L recently, either directly or indirectly. The direct prediction refers to a ML model or several models in sequence with κ_L as the final output. The majority of studies over the years fall under this category. Several ML models have been trained on 110 half-Heusler compounds by compiling elemental, compound, and compound-elemental descriptors, obtaining high validation accuracy and revealing the bond distance as the most important descriptor for κ_L [184]. Diamond-like materials were explored for ultra-high κ_L using a small training datasets through transfer learning of the three-phonon scattering channel volume, or “P_3” for short [118]. On the other hand, the indirect prediction approach first predicts lower-level properties eventually leading to κ_L, including the atomic forces and the IFCs which are required for κ_L. To date, little research has been seen in this subarea. Notably, providing lower-level physics introduces several advantages over the direct method. Firstly, with the atomic forces and/or IFCs, one can compute the full

phonon properties, e.g., phonon dispersions, temperature-dependent κ_L, and scattering rates. This allows for in-depth study of the phonon properties at the high-throughput level without needing to rely strictly on pure DFT. The temperature dependent κ_L is especially desirable for high-temperature applications whereas most studies involving ML focus only on room temperature. Secondly, because κ_L is not directly computed, variables involved in the BTE calculation may be modulated, such as the inclusion of higher order anharmonicity beyond 3rd order and off-diagonal contribution. Thirdly, atomic forces and IFCs are much more abundant than single κ_L values associated with each structure, allowing for improved training and potential transferability of information among diverse structures.

Here, we demonstrate a “bottom-up” approach by developing the elemental and spatial descriptors of the Elemental Spatial Density Neural Network Force Field (Elemental-SDNNFF). We evaluated 77,091 cubic structures of varying space-groups and chemical species. This is made possible by our model that provides sufficient flexibility to distinguish many unique atomic environments for the high-throughput calculation of full phonon properties. Because the forces are provided, thermodynamic stability of these structures can be determined *via* their predicted phonon dispersions which is not possible in direct methods. Here, the model is initially trained on a small subset of 3,107 structures and is iteratively improved on a larger dataset of 77,091 structures with active learning. Data augmentation is incorporated whereby equivalent atomic environments are rotated to provide a ~3× boost to the total atomic forces

for training. Then, the final model is deployed to predict the complete phonon properties of the remaining stable structures, with speed of three orders of magnitude faster than full DFT calculations. Our model is capable of growing and being adapted to new datasets for exploration of even larger unseen materials.

5.2 Structure Generation and Active Learning

To develop an ML model for evaluating atomic forces, it is crucial to prepare a sufficiently large and diverse dataset of various atomic environments. In theory, there are limitless combinations of atomic environments considering structural symmetries, chemistries, and displacements that one may encounter during the evaluation of atomic forces. This is especially true when involving “random” or stochastic displacement methods for generating IFCs such as compressive sensing lattice dynamics (CSLD) [138]. As ML methods specifically supervised learning are interpolative by nature, selection of training or reference data for evaluation of seemingly infinite possible atomic environments requires a human-free or self-informative approach for efficient model training from costly DFT calculations. The crystalline structures analyzed here are borrowed from the OQMD database [134,182] and are categorized into 16 prototypes spanning four cubic space groups as seen in Table 5.1. Some structures are split from the pool for initial training and evaluated with DFT for their atomic forces. From DFT, the IFCs and subsequent phonon properties are also gathered for comparison. Accordingly, the rest of the structures without DFT are marked as ‘unexplored’ left for the model evaluation stage.

Table 5.1 Distribution of all cubic structures used for training and screening.

Space group no.	Structure type	Structures calculated by DFT	Predicted stable structures by DFT	Unexplored structures	Predicted stable unexplored structures
216	AB	61	58 (96.7%)	209	24 (11.5%)
	ABC	452	446 (98.9%)	3,144	2,980 (94.8%)
	ABC_2	134	127 (95.5%)	5,901	2,324 (39.4%)
	ABCD	553	549 (99.3%)	8,059	5,825 (72.3%)
221	AB	365	359 (98.1%)	786	323 (41.1%)
	AB3	338	337 (99.7%)	1,695	1,059 (62.5%)
	AB_3C_3	51	47 (92.2%)	785	64 (8.2%)
	AB_3C_{12}	5	3 (60.0%)	218	18 (8.3%)
	ABC_3	103	100 (96.2%)	597	119 (19.9%)
225	AB	163	156 (95.7%)	518	145 (28.0%)
	AB_3	160	146 (91.2%)	1,705	525 (30.8%)
	ABC_2	581	576 (99.0%)	12,643	7,472 (59.1%)
	ABC_2D_6	92	86 (92.5%)	39171	5,227 (13.3%)
227	AB2	52	48 (94.2%)	282	84 (29.8%)
	AB_2C_4	68	66 (98.5%)	1,298	864 (66.6%)
	ABC_2	4	3 (75.0%)	80	6 (7.5%)
Total		3,182	3,107 (97.7%)	77,091	27,059 (35.1%)

*Values in parenthesis indicate percentage out of the left column.

We first perform structure filtration by elements, formation energy, and energy above hull, and then perform structure optimization observed by the blue path in Figure 5.1. Arrow and box colors represent different regimes of the workflow with blue, orange, red, and green representing the structure generation phase, initial model training phase, iterative model training or active learning phase, and the application or deployment phase, respectively. In the application phase, the final ML model is applied to evaluation of atomic forces, based on which the interatomic force constants are fitted and phonon properties are subsequently predicted. The purpose of filtration is to reduce the structure count and increase the likelihood of stability in subsequent DFT calculations of phonon dispersions. After structure optimization, we replicate primitive cells into supercells and displace all atoms by fixed 0.03 Å in random directions to create diversity in the atomic environments and facilitate training, which is also standard for IFC calculations. Additionally, for IFC fitting with CSLD, such introduction of random displacements helps mitigate the innate poor energy conservation of direct force field models, as demonstrated in previous works [96,185]. Specifically, due to the nature of direct force prediction, IFC fitting with methods such as finite difference method (FDM) does not guarantee zero or near-zero forces for atoms in equilibrium. Such atoms are abundant in FDM whereby only one or two atoms are displaced for IFCs up to the third order, causing significant disagreement in the force-sensitive lattice thermal conductivity. Thus, stochastic methods such as CSLD mitigate the energy conservation issue by displacing all

atoms, generating a “noise-cancelling” effect for subsequent IFCs fitted by predicted forces.

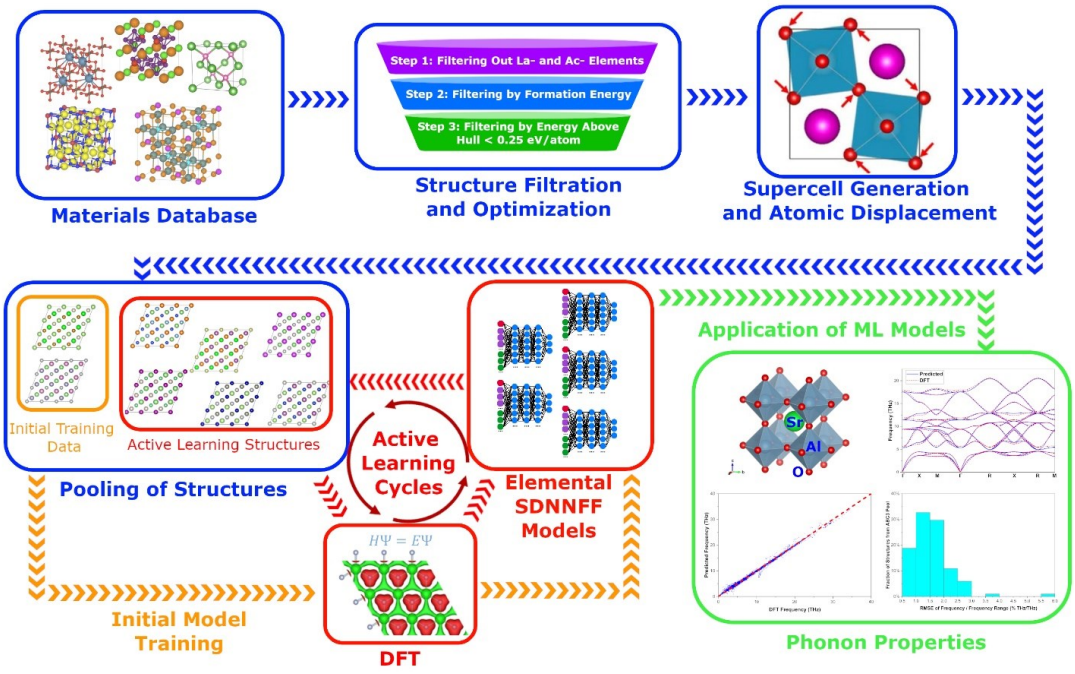

Figure 5.1 Workflow for training Elemental-SDNNFF models.

Afterwards, we randomly select a small fraction of displaced supercell structures to serve as the initial training data for the models. This is passed to DFT calculations, and the resulting ground truth atomic forces and local atomic environment are passed to five initialized models with similar architecture but different weights. Here, we used a small subset of existing DFT data previously calculated for phonon properties. Although we do not introduce techniques to choose the initial structures for training, we recommend those such as the principal component analysis (PCA) shown to improve model representation especially for out-of-trend structures [186]. After the DFT step, we perform data augmentation in which atomic environments are rotated according to nearest

neighbor rules, generating approximately 2-3× increase in the existing dataset. This enhances the dataset diversity for model training in addition to the already abundant $N \times t$ dataset, where N is the number of atoms per simulation (or per supercell) and t is the number of DFT simulations. As an aside, future work is planned to introduce rotational equivariance to the existing Elemental-SDNNFF model for automatic consideration of rotated atomic environments without the need for said nearest neighbor rules and rotation matrices [187]. This should drastically improve the training efficiency by reducing the training size while maintaining the current rotational covariance of the force field.

Taking advantage of the interpolative nature of neural networks, these models serve as a “committee” that will judge the remaining untrained or active learning structures for atomic environments. Structures yielding high uncertainty in the forces indicates poor representation of the corresponding atomic environments and DFT forces. Using Equation 4.10, we set the uncertainty threshold to $\varepsilon_i > 50$ meV/Å well above the force error of the model to guarantee poorly represented structures in the committee. Those structures with uncertainty above the threshold are then passed to further DFT calculations and retrained into the model, forming a closed loop with iterative self-improvement (red circled arrows in Figure 5.1). Once the number of recommended structures converges to near-constant value, we then publish the model for evaluation of atomic forces and subsequent phonon properties for all unseen structures (green path in Figure 5.1). After seven rounds of active learning with combined data augmentation, 29.4 million atomic environments are successfully trained into the network.

5.3 Network Performance and Prediction of Phonon Properties

To benchmark the performance of the trained model for phonon properties prediction, we first examine the atomic force accuracy on a small subset of 400 untrained structures as seen in Figure 5.2. The structures shown here span half, full, quaternary Heusler structures, and single and double perovskites. We compare the performance of the model to CHGNet, which was recently proposed as a universal potential energy surface model [188]. We found a force root mean square error (RMSE) of 29.3 meV/Å and 121 meV/Å for the Elemental-SDNNFF and CHGNet models, respectively, showcasing the competitive performance of our model and its consistency remaining close to the training RMSE of 18.6 meV/Å.

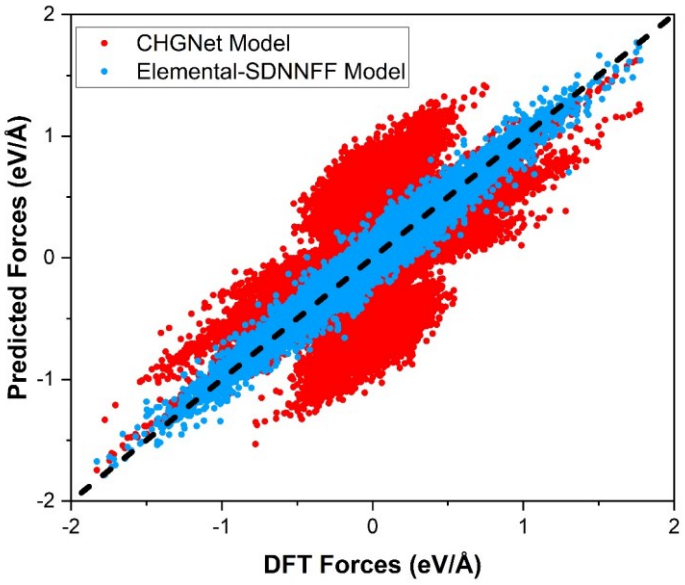

Figure 5.2 Comparison of the atomic forces in 400 untrained structures between predicted and target values.

Thereafter, the errors relative to DFT for the phonon dispersion and corresponding κ_L at 300K are shown in Figure 5.3. In Figure 5.3.a, the RMSE of the frequency is divided by the frequency range of the corresponding dispersions to normalize and merge the data to a single histogram and is shown as a percentage. The insets show phonon dispersions linked to the relative error containing DFT (black lines) and prediction (red lines) for visualization. The average error is 1.88% which is excellent as seen by the insets of sample dispersions in Figure 5.3.a. In Figure 5.3.b, κ_L at 300 K between DFT and the developed single neural network model are compared for 3,107 stable structures. In Figure 5.3.b, the log value of the DFT and predicted κ_L yields an R^2 of 0.89 with a mean average error (MAE) of 0.254 log($Wm^{-1}K^{-1}$), meaning that the predicted κ_L is on average within 1.795 times the DFT value and is shown by the structures within the dashed lines representing two times the perfect agreement. The prediction capability is competitive with the 0.12 MAE and 0.87 R^2 error presented by the random forest model trained on ~10^3 materials [121]. Additionally, 103 untrained structures with κ_L from DFT are evaluated by our model for validation and are compared in the inset of Figure 5.3.a. Out of these structures, 67 were predicted owning < 1 $Wm^{-1}K^{-1}$ and 36 remain within the same range from DFT values. Notably, at the lower end of the κ_L range, the model tends to underpredict the κ_L with greater intensity approaching the ultralow range of predicted 0.1 $Wm^{-1}K^{-1}$. This is due to the highly sensitive nature of the phonon transport toward the extrema of the κ_L, specifically from the quality of atomic forces in displaced supercells and eventually the 3rd order IFCs [96].

Nonetheless, the materials with prediction under 0.1 $Wm^{-1}K^{-1}$ are likely to remain within 1 $Wm^{-1}K^{-1}$ range and our model is effective at filtering candidates with ultralow κ_L.

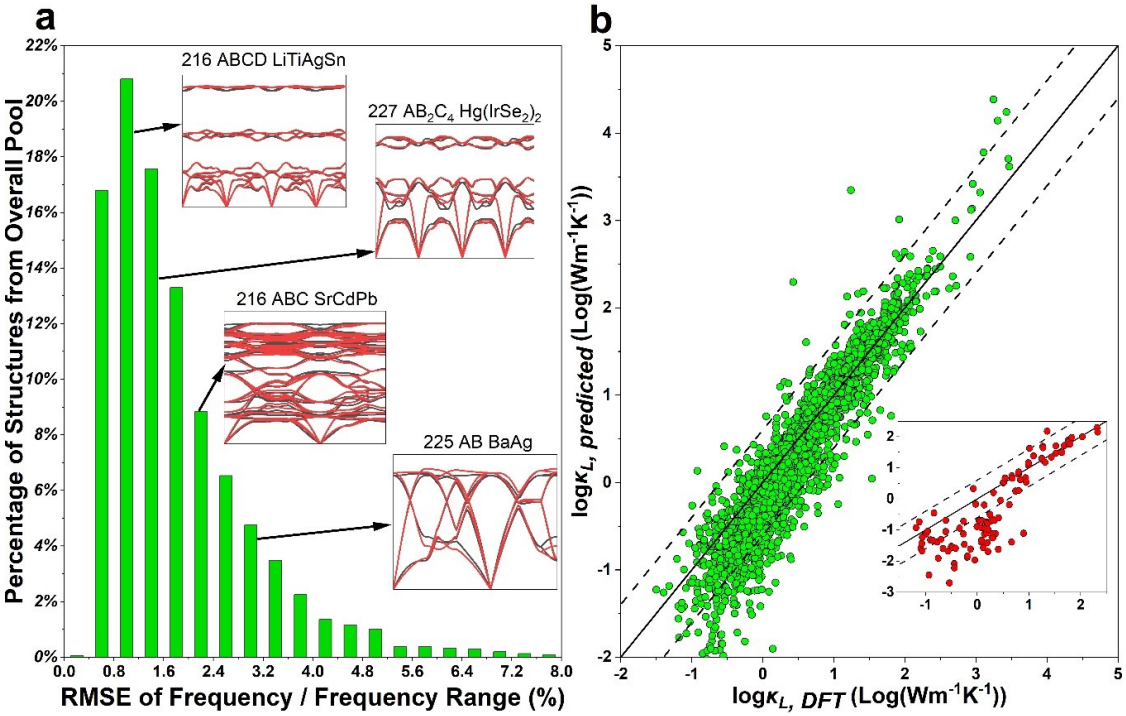

Figure 5.3 Comparison of phonon properties for 3,107 structures between the model and density functional theory

The advantage of our *bottom-up* ML approach for phonons manifests from the plethora of information from standard phonon calculation packages when provided a set of predicted atomic forces. Indeed, by default phonon frequencies and scattering matrix elements required for iterative BTE are computed in advance. From the phonon frequencies, information like the speed of sound, constant volume heat capacity, and the three-phonon scattering phase space may be readily computed. Here, we compare these properties from our neural network model with those from DFT calculations to further understand more

about the underlying contributions to the predicted κ_L. Indeed, the constant volume heat capacity (c_v) is directly involved in computing κ_L along with the phonon group velocities and scattering rates. Additionally, the speed of sound (V_s) is a partial representative of the group velocities for long-wavelength acoustic modes in crystals [2]. The three-phonon scattering phase space (P_3), is a quantitative measurement of the number of three-phonon scattering channels. Unlike c_v and V_s, higher P_3 is indicative of larger scattering rates and thus lower κ_L [118]. However, akin to c_v and V_s, P_3 requires only the second order IFCs and therefore is computationally inexpensive as a result after the model is trained. Finally, the mean square displacement (MSD) of vibrating atoms, usually observed in finite temperature molecular dynamics, may also be computed from the phonon frequencies and eigenmodes [47]. To describe structures with a single value, only the maximum MSD, shortened to max MSD, among all atoms in the primitive cell is assigned.

In Figure 5.4, the comparison of c_v, V_s, P_3, and maximum MSD between the neural network model and DFT is shown with the corresponding R^2 and MAE values. Exceptional agreement is found for c_v, followed by V_s, P_3, and log(max MSD). With the small error in phonon frequencies, the mode-weighted global property c_v requires the mode-dependent phonon frequency as a direct input and is summed up over a dense sampling of Brillouin zone, and thus expectedly owns the best agreement with DFT [48]. The max MSD is also constructed similarly and thus owns high accuracy with DFT. Although V_s is also directly computed from dispersions, the gradient of phonon frequency with respect to

wave vector is required and is consequentially more sensitive to the predicted atomic forces than c_v. Finally, P_3 also uses the phonon frequencies directly but involves a counting of three phonon collisions by energy and momentum conservation. In other words, the error propagated from atomic forces into the phonon frequencies is compounded resulting in the largest error out of the other three harmonic phonon properties. Interestingly, while Figures 5.4.a-5.4.c for the most part experience an even spread of error, the scatter plot for the maximum MSD in Figure 5.4.d is shown with increasing disagreement at higher maximum MSD. This is because higher maximum MSD, corresponding to softer phonon modes, usually indicates lower κ_L, and the second order IFCs are more sensitive to the atomic displacements and corresponding forces. In such a case, we anticipate that increasing the atomic displacements can better capture the anharmonicity and hence the potential energy surfaces.

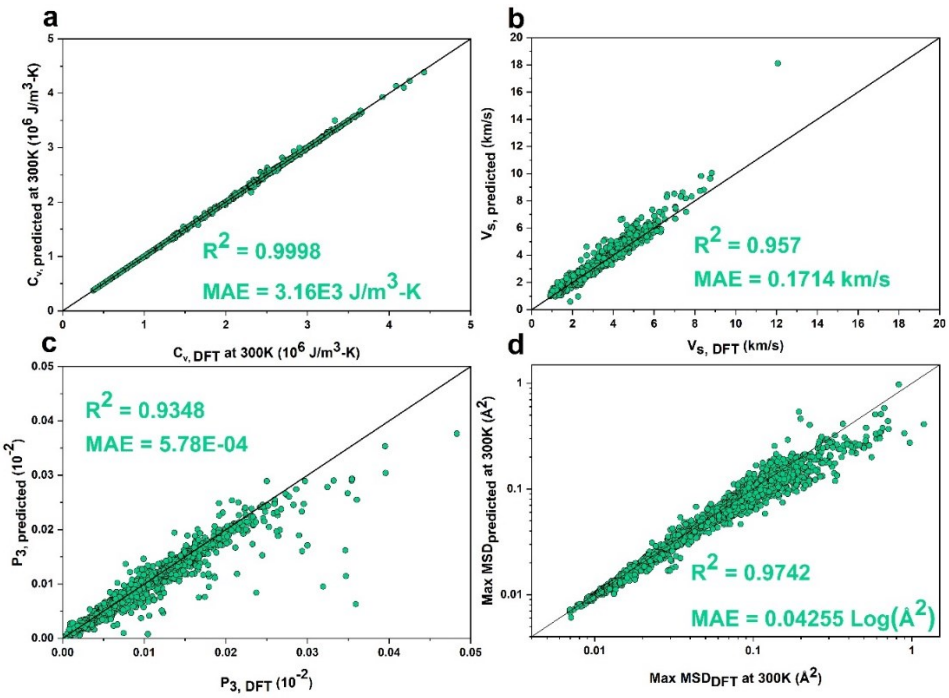

Figure 5.4 Comparison of the harmonic properties between the model and density functional theory.

5.4 Quantification of Ultralow Lattice Thermal Conductivity with Predicted Properties

Out of the 77,091 cubic materials set aside for evaluation, 27,059 structures are predicted by our trained Elemental-SDNNFF model to have no imaginary phonon frequencies, and thus being potentially thermodynamically stable. These structures are then evaluated for predicting full phonon properties including c_v, V_s, P_3, max MSD, and κ_L. In the previous section, we show how the trained model represents DFT-level IFCs while maintaining a speed-up on the order of ~10^3 for systems with 100 atoms, as seen in Figure 5.5. This was made possible due to the initial training set and subsequent active learning iterations

generating millions of data points to best fine-tune the model for handling many structures and chemistries.

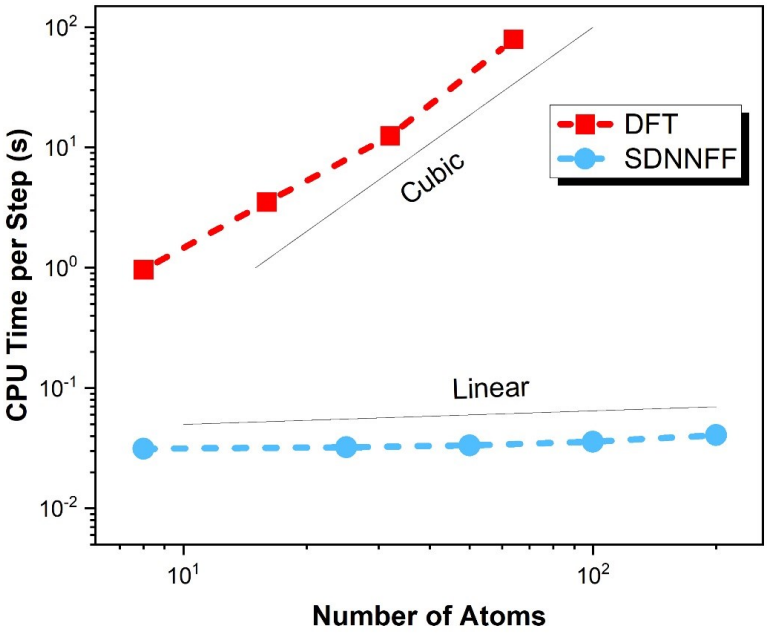

Figure 5.5 The CPU time per step against the number of atoms in a system from density functional theory and the model.

Despite replacing DFT with a machine learned model, the process of computing anharmonic IFCs and subsequently iteratively solving BTE is still time consuming for several tens of thousands of hypothetical structures. Therefore, given the large data of phonon properties evaluated by our model, quantification of trends for κ_L with structural and harmonic properties is desired to search materials with known thermodynamic stability. Previous studies suggest that κ_L is strongly correlated with several physical parameters, including volume of the

unit cell V_{cell} [121], specific heat capacity c_v [128], sound velocity V_s [117], three-phonon scattering phase space volume P_3 [118], and thermal max MSD [189]. Henceforth, we have experimented with linear combinations of c_v, V_s, P_3, max MSD, and V_{cell} to correlate with κ_L. We found the max MSD by itself has the best performance as a descriptor for κ_L of crystals. The reason is most likely due to the major contribution of harmonic phonons in thermal max MSD for κ_L when compared to those listed above. Additionally, MSD may be computed as a function of temperature and is more useful to observe temperature-dependent trends. A generally inverse-linear relationship is observed between the logκ_L and the log(max MSD) in Figure 5.6. Note, log(max MSD) is normalized here in $\mathbb{R} \in$ [0,1] based on values found from the DFT set for ease of comparison. Figure 5.6.a provides evidence of linearity through comparison of κ_L and the max MSD. The fitted red line shows a decreasing trend of κ_L with increasing max MSD. Structures with extremely high maximum MSD are indicative of rattling atoms, in which strong phonon-phonon scattering and ultralow κ_L is prevalent [189]. Given the agreement of the max MSD between DFT and predictions in Figure 5.4.d, Figure 5.6.b demonstrates the prediction of both κ_L and max MSD for the 25,901 unexplored structures out of the stable 27,059 pool, since some structures failed in BTE calculations and κ_L was not plotted. Again, the trend remains inversely proportional to the descriptor. We do notice that the newly fitted blue line shows a steeper slope in comparison to the previous red line by DFT (also shown in the same plot for comparison), and the difference between the two lines deviates with increasing maximum MSD. The most probable cause is the underprediction

of κ_L at the lower extreme (Figure 5.3.b) and increased error in the MSD at the higher extreme (Figure 5.4.d). Although the ultralow κ_L may be underestimated on the log scale, these predictions remain highly beneficial for quickly marking structures with potential as thermal insulators. To quickly filter ultralow κ_L materials, a max MSD is set such that the value of the fitted line is 1.795 $Wm^{-1}K$. This is chosen deliberately knowing predicted values of κ_L are within 1.795 times the DFT value which aids the later filtration for structures less than 1 $Wm^{-1}K^{-1}$. As such, the maximum MSD filter is set to 0.076 $Å^2$ or 0.464 on the normalized log(max MSD) plot. Specifically, we found 9,306 total structures with normalized log(max MSD) higher than 0.464. Out of these structures, the κ_L of 8,873 (95.4%) structures are less than 1 $Wm^{-1}K^{-1}$. For normalized log(max MSD) less than 0.464, out of 16,596 structures, the κ_L of 4,590 are less than 1 $Wm^{-1}K^{-1}$. This means, the success rate for filtering structures is 66% (8,873 out of 13,461 structures) for those with κ_L less than 1 $Wm^{-1}K^{-1}$. Thus, the maximum MSD is a reliable descriptor for indicating highly unique structures with out-of-trend values of atomic displacement and corresponding ultralow κ_L. Such a critical value of maximum MSD 0.076 $Å^2$ may be used in future works to identify potential candidates for thermal insulators in cubic structures.

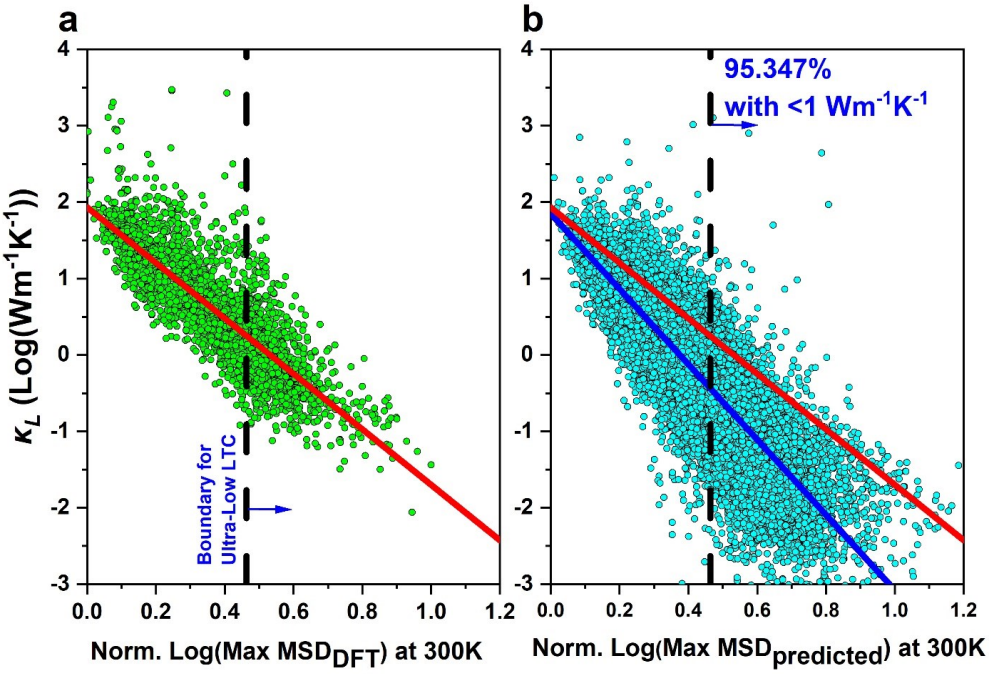

Figure 5.6 Plots for the lattice thermal conductivity against maximum thermal mean square displacements.

Rattling effect has been proved to induce large MSD in many systems. To generalize structures with high probability for rattler atoms, we plot the average of the MSD for each element across the 25,901 pool in Figure 5.7. The color bar represents the magnitude of the average MSD in $Å^2$. Hydrogen and all alkali metal elements, including Li, Na, K, Rb, and Cs, have the highest average MSD among all 63 elements covered here. Some alkaline earth metal elements including Mg, Ca, Sr, and Ba have medium MSD. Interestingly, halogen elements including Cl, Br, and I also possess high MSD. N stands out among the nonmetallic elements with the next largest being C, P, and Si. Tl also stands out as a semimetallic element next to Pb and Sn with significantly lower MSD.

Metallic elements including Hg, Cd, Ag, and Au are the highest in their category although their MSD is much lower than their alkali metal counterparts.

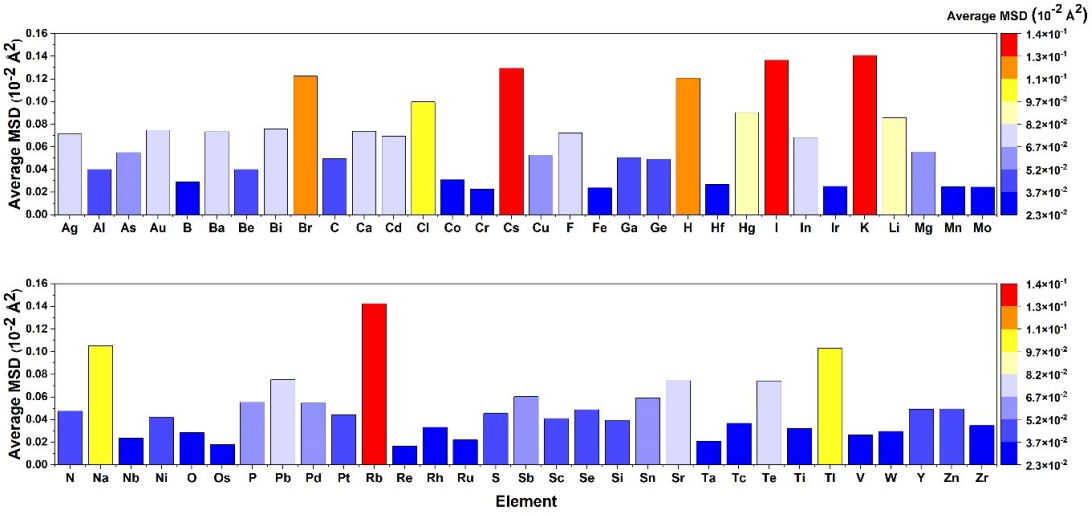

Figure 5.7 Average thermal mean square displacement at 300K across 25,901 predicted stable structures.

To visualize the spread of predicted κ_L in the dataset, Figure 5.8 displays the t-distributed stochastic neighbor embedding (t-SNE) using Elemental-SDNNFF structure input vectors. For simplification, only one point per structure is implemented corresponding to a single Elemental-SDNNFF vector centered at the unit cell rather than on a per-atom basis. In Figure 5.8.a, the distribution of space group indicates an overlap of structures with space group number 216 and 225 whereas a majority of structures with space group number 221 and 227 form visible clusters. This is sensible given space group 216 is different from space group 225 by just a vacant lattice site (ABC vs. ABC_2) or a different element on the same lattice site (ABCD vs. ABC_2). Figure 5.8.b focuses on the predicted κ_L

where observable regions of thermally insulating materials (blue) are highly contrasted from thermally conductive materials (red). Mainly, the upper left of the figure contains a mixture of space group 216 and 225 structures with ultralow κ_L with some additional blue regions along the bottom and top outer edge mostly corresponding to 221 and 225. This is further manifested in Figure 5.8.c where regions are instead highlighted by the predicted normalized log(max MSD). To highlight the relationship between κ_L and max MSD, we subtract the value from unity to match the properties based on color. As seen by the comparison between Figures 5.8.b and 5.8.c, both figures form identical structure mappings of the predicted κ_L and normalized log(max MSD) values, supporting their strong correlation. Additionally, the congregation of certain out-of-trend structures with extremely high or low κ_L indicates a correlation between Elemental-SDNNFF input vector and phonon conductors and insulators, suggesting strong structural-property relationship with phonon transport. Overall, the t-SNE plots encompass the wide range of unique structures and physics manifested when applying machine learned atomic force fields such as the Elemental-SDNNFF.

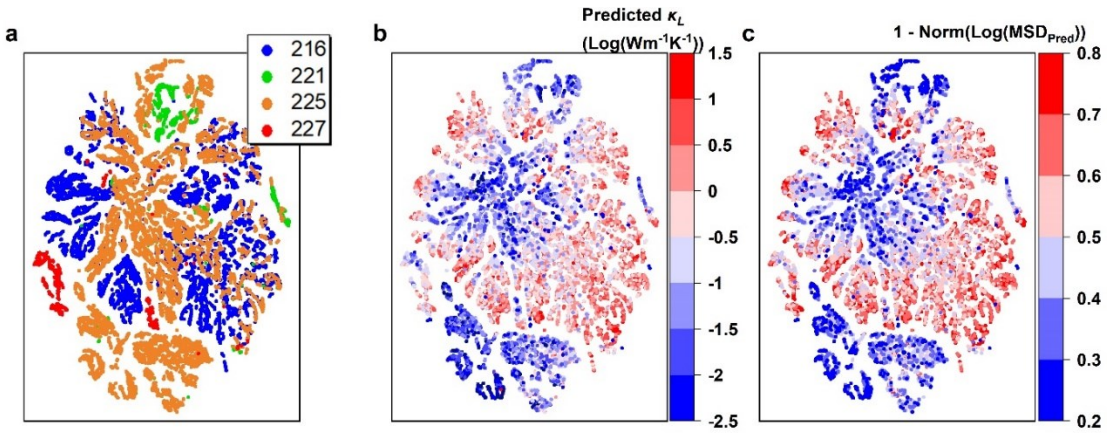

Figure 5.8 t-SNE plots of model input vectors for 25,901 structures.

5.5 Insight from Bonding and Anti-Bonding

In the previous section, we related thermal displacements to the κ_L but do not discuss the effect of chemical bonding. Here, we further analyze our predicted structures with Crystal Orbital Hamilton Population (COHP) [190] to quantify the contributions to the bonding and antibonding states. To assign single values of bonding and antibonding to each structure, we perform integration over COHP curves for each atomic pair as evaluated by LOBSTER code and take the average [191]. Figure 5.9 displays the resulting bonding and antibonding with highlighted log values of predicted κ_L to observe trends. Notably, at low bonding values (e.g., < 200) and high antibonding (e.g., > 1), only κ_L < 3 Wm^{-1}K^{-1} exists. This region contains low interatomic bonding strength and high phonon anharmonicity, resulting in ultralow κ_L. Our observed trend of high antibonding indicates strong phonon anharmonicity and is consistent with recent studies by full DFT calculations on other systems [116,176,192,193]. On the other hand, the

high bonding (> 200) region seems to contain all ranges of κ_L. This is understandable from the physics point of view, whereby the κ_L is governed by two major mechanisms of interatomic bonding strength and phonon anharmonicity and thus the bonding/antibonding contributions might be competing. It is also worth pointing out that, the application of COHP is a low-cost indicator of κ_L requiring only the DFT calculations on primitive cells after structure optimization, which is very promising for the filtration of structures with anomalously low κ_L.

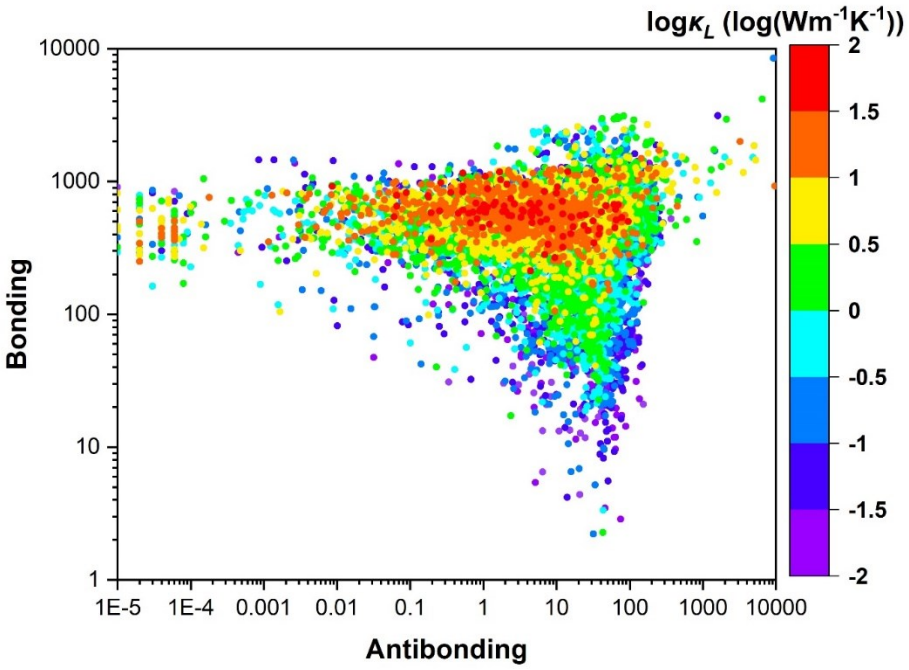

Figure 5.9 Crystal orbital Hamilton population analysis for 13,718 stable structures from the prediction pool.

5.6 Off-Diagonal Thermal Transport Analysis

Recently two different mechanisms for phonon transport in solids have been discussed [194,195]. In crystalline materials, heat carriers propagate and scatter in a particle-like behavior as described by Peierls-Boltzmann transport picture for phonon wave-packet dynamics. Such populations have a well-defined energy (frequency) and therefore can be interpreted as particle-like excitations with a well-defined wave vector ($\boldsymbol{q}$) and mode index (s), and corresponding lattice thermal conductivity is denoted as κ_P. In contrast, in glass materials, heat carriers behave wave-like, hopping via a Zener-like tunneling between quasi-degenerate vibrational eigenstates, as described by the Allen-Feldman theory. Such coherences do not have an absolute energy and cannot be related to a single eigenstate. Rather, they describe oscillations between pairs of eigenstates and correspond to an evolution which does not preserve the nature of the single-particle excitation. Very recently, the importance of such coherences' contribution to overall lattice thermal conductivity describing the wave-like interband tunneling of phonons, dubbed as two-channel thermal transport or off-diagonal contribution of heat-flux operator, has been theoretically formulated and experimentally validated in ultralow κ_P materials [123,196,197]. With the phonon property of large-scale crystals being predicted herein, it is intuitive to explore the two-channel thermal transport behavior of these materials.

The contributions of off-diagonal components ($\kappa_c^{\alpha\beta}$) to the total thermal conductivity is obtained by [123,198]

$$\kappa_c^{\alpha\beta}=\frac{\hbar^2}{k_B T^2 N\Omega}\sum_{q}\sum_{j,j'}^{j\neq j'}\frac{\omega_j(\mathbf{q})+\omega_{j'}(\mathbf{q})}{2}V_{j,j'}^{\alpha}(\boldsymbol{q})V_{j,j'}^{\beta}(\boldsymbol{q})\times \tag{5.1}$$

$$\frac{\omega_j(\mathbf{q})n_j(\mathbf{q})\left(n_j(\mathbf{q})+1\right)+\omega_{j'}(\mathbf{q})n_{j'}(\mathbf{q})\left(n_{j'}(\mathbf{q})+1\right)}{4\left(\omega_j(\mathbf{q})-\omega_{j'}(\mathbf{q})\right)^2+\left(\Gamma_j(\mathbf{q})+\Gamma_{j'}(\mathbf{q})\right)^2}\times\left(\Gamma_j(\mathbf{q})+\Gamma_{j'}(\mathbf{q})\right)$$

where $\hbar$, k_B, T, N, and Ω are the reduced Planck constant, Boltzmann constant, absolute temperature, the number of q-points sampled in reciprocal space, and volume of the unit cell, respectively. $\omega_j(\mathbf{q})$, $\Gamma_j(\mathbf{q})$, and $n_j(\mathbf{q})$ are the frequency, linewidth, and the equilibrium Bose-Einstein distribution function with wave vector $\boldsymbol{q}$ and branch j, respectively. $V_{j,j'}^{\alpha}(\boldsymbol{q})$ is the off-diagonal elements ($j\neq j'$) of velocity matrix with direction α, and can be calculated by

$$V_{j,j'}^{\alpha}(\boldsymbol{q})=\frac{1}{\omega_j(\mathbf{q})+\omega_{j'}(\mathbf{q})}\langle \mathrm{e}_j(\mathbf{q})\left|\nabla_q^{\alpha}\mathbf{D}(\mathbf{q})\right|\mathrm{e}_{j'}(\mathbf{q})\rangle \tag{5.2}$$

where $\mathrm{e}_j(\mathbf{q})$ and **D** are the eigenvector and dynamical matrix, respectively. In our work, we modified the original ShengBTE code [139] to output the off-diagonal elements of velocity matrix and then calculate the off-diagonal thermal conductivity based on Equation 5.1.

For calculations of lattice thermal conductivity contribution by coherence phonons (κ_c), 1,000 structures from each pool of materials with low (< 1 $Wm^{-1}K^{-1}$), medium (1 – 10 $Wm^{-1}K^{-1}$), and high (> 10 $Wm^{-1}K^{-1}$) κ_P by traditional BTE solution are randomly chosen with the same parameters and force constants from the model prediction step. From these structures, 869 low, 995 medium, and

999 high κ_P BTE calculations successfully converged with average off-diagonal ratios κ_c/κ_P of 2.59, 5.02×10^{-2}, and 4.35×10^{-3}, respectively. The κ_c/κ_P show good agreement between DFT results and prediction by our Elemental-SDNNFF model (Figure 5.10).

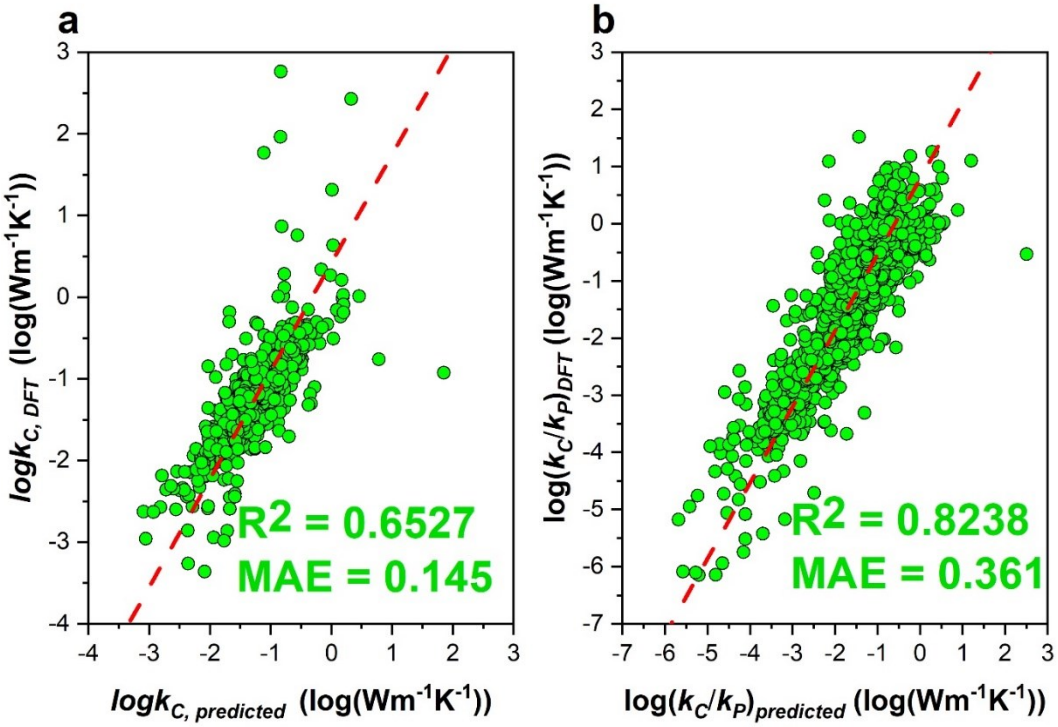

Figure 5.10 The off-diagonal contribution to lattice thermal conductivity computed for 2,398 structures.

In Figure 5.11, we compute the percentage of off-diagonal contribution which is defined as κ_c/κ_{total} where $\kappa_{total} = \kappa_c+\kappa_P$. The agreement between our Elemental-SDNNFF predictions on selected 2,863 unseen data and DFT from 2,397 training data is indicated by their overlap and matching linear trend. The dashed black line indicates 50% from off-diagonal contribution, above which the wave-like heat conduction is dominated. Good agreement between our Elemental-SDNNFF

predictions on unseen data and DFT results from training data is observed. Strong linear-like correlation trend is found between percentage of κ_c from κ_{total}. Generally speaking, the lower κ_P is, the higher percentage of off-diagonal contribution by κ_c could have. For extremely low κ_P materials, e.g., $\kappa_P \sim 0.1$ $Wm^{-1}K^{-1}$, the κ_c could contribute as high as 50% or even 70% of κ_{total}, showing dominant role of contributions of the coherences even at room temperature. Similar phenomenon is also found in perovskite $CsPbBr_3$ at room temperature and $La_2Zr_2O_7$ at high temperature [123,197]. We also noticed that, for the same κ_{total}, the κ_c contribution can differ quite largely among different structures, leading to a very wide spread of the log-scale percentage contribution in Figure 5.11. This implies that, the precise off-diagonal contribution percentage depends on detailed phonon band structures and mode level phonon anharmonicity of different materials, rather than the single absolute value of κ_P. Figure 5.11 also shows that there are considerable amount of structures with dominant wave-like heat conduction, instead of particle-like conduction as predicted by traditional Peierls BTE. Our model clearly determines the crossover from particle-like to wave-like heat conduction (dashed line in Figure 5.11).

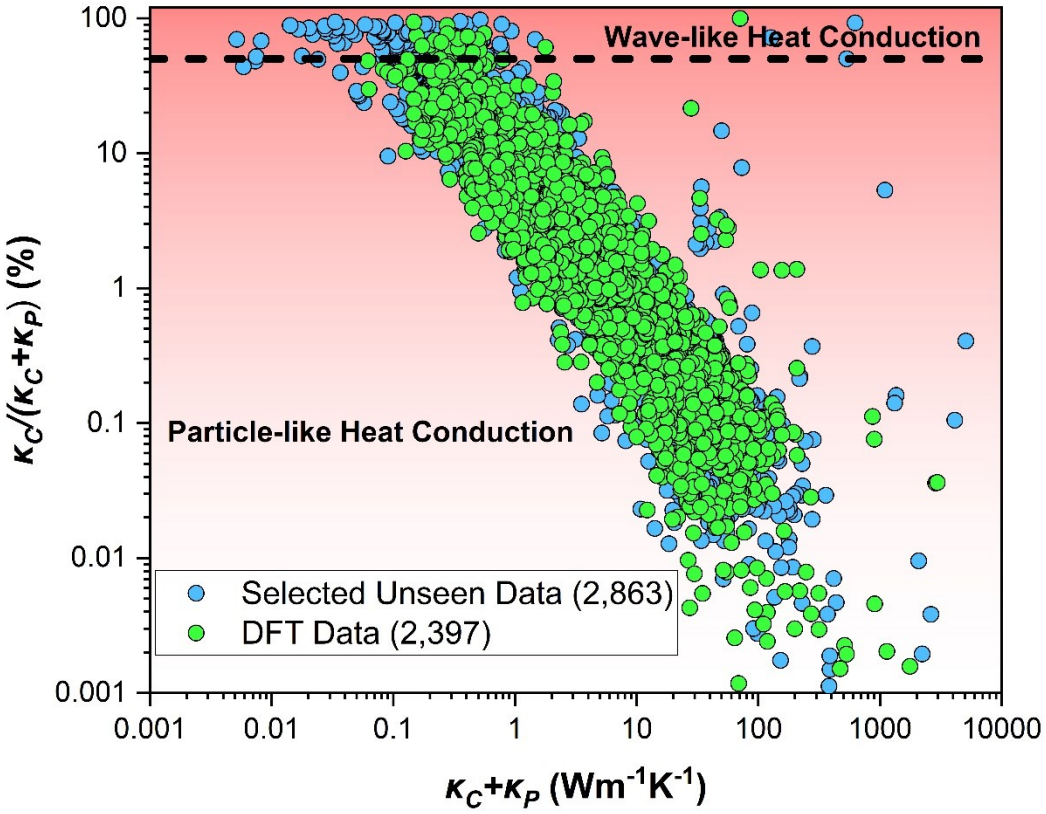

Figure 5.11 The off-diagonal contribution against the total lattice thermal conductivity at 300K.

To observe mode-level contributions of κ_P and κ_c, three low κ_P materials (NaKAs $KIrCs_2Cl_6$, and $CuPdSr_2$ with space group no. 216, 225, and 216, respectively) from the untrained pool are chosen with varying coherence contribution for comparison. The frequency-dependent relaxation time is plotted in Figure 5.12. The size of circles represents the magnitude of κ_P, whereas the color represents the ratio of κ_c/κ_{total}, i.e., the off-diagonal contribution to overall thermal transport. The diameter of circles is scaled up equally between predicted

and DFT plots for viewing. The materials are provided in order of increasing κ_P with (prediction value, DFT validation value) of (0.037, 0.075), (0.101, 0.127), and (0.68, 0.935) $Wm^{-1}K^{-1}$, respectively. The κ_c values for (predicted, DFT) results are (0.222, 0.205), (0.252, 0.226), and (0.07289, 0.08282) $Wm^{-1}K^{-1}$, respectively. From comparison between the predicted (left column in Figure 5.12) and DFT (right column in Figure 5.12) results, the general trends of relaxation time against frequency are captured, confirming again the accuracy of our ML model. Despite this, κ_P contributions from the low frequency range are more pronounced in size and density by eye in the DFT results (Figures 5.12.b, 5.12.d, and 5.12.f) than in the prediction results (Figures 5.12.a, 5.12.c, and 5.12.e). This agrees with the previously observed underprediction of κ_P for structures with $\kappa_P <$ 1 $Wm^{-1}K^{-1}$. In addition, the κ_c of NaKAs and $KIrCs_2Cl_6$ share similar values and is pronounced by the presence of high mode-level κ_c contribution in the entire frequency range. For $CuPdSr_2$, a vast majority of points own low contribution by coherence phonons thus reducing the κ_c. From observation, κ_c contributions are overshadowed by dominating mode-level κ_P throughout the entire frequency range, leading to lower κ_c/κ_{total} ratios as previously seen in Figure 5.11.

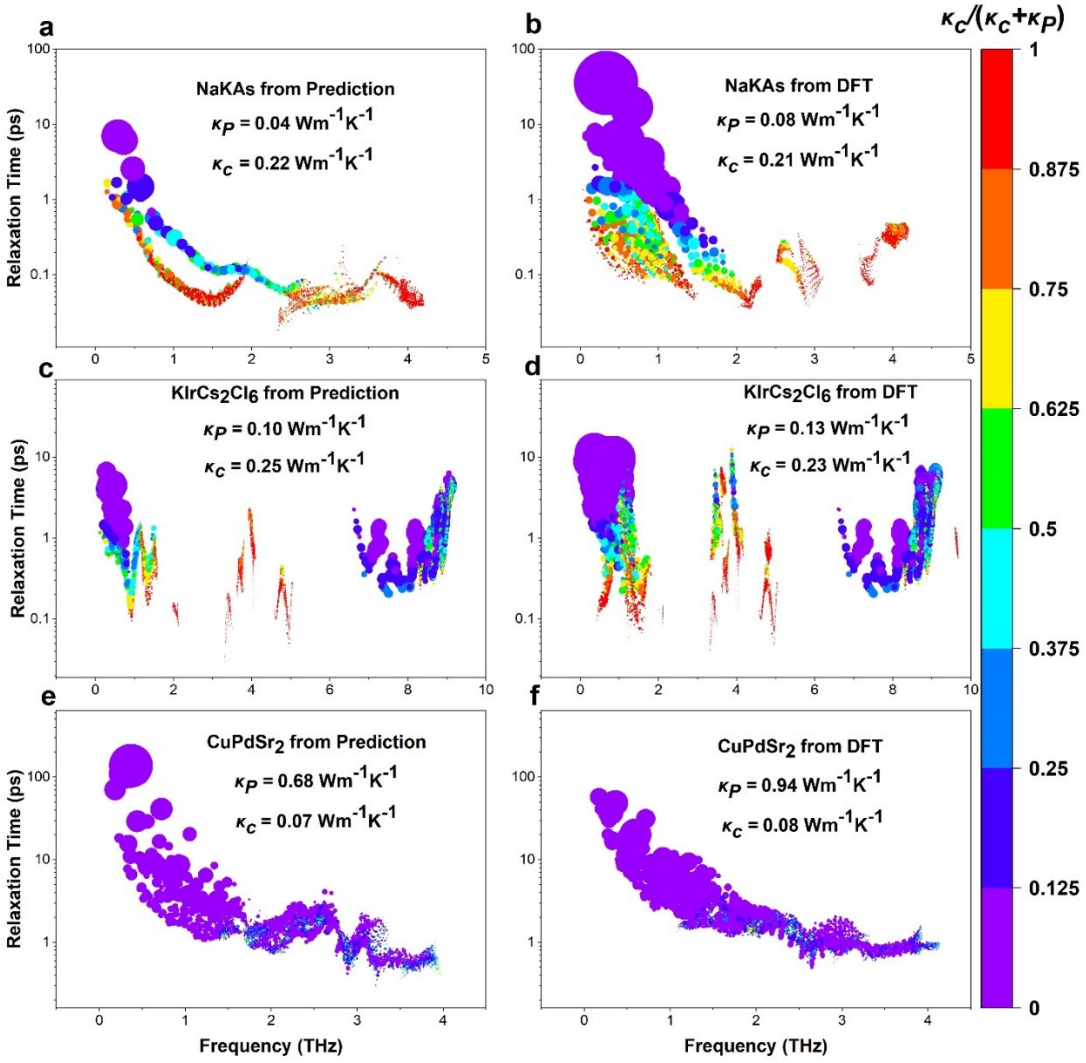

Figure 5.12 The frequency dependent relaxation time for NaKAs, $KIrCs_2Cl_6$, and $CuPdSr_2$.

5.7 Conclusion

We demonstrate the development of a new “bottom-up” machine learning approach through accurately and efficiently predicting comprehensive phonon properties of ~80,000 crystalline materials, which is realized by evaluating the atomic forces with sufficient DFT representation. The “query-by-committee”

active learning scheme allows iterative improvement of the models by simultaneous prediction of atomic forces in the unseen pool of displaced structures. The final model is deployed for constructing IFCs for both observed and unseen structures. Given the good accuracy of phonon properties such as vibrational frequencies and κ_L, we exploit the abundant 25,901 pool of structures that are predicted to be thermodynamically stable by our model and quantify thermal insulators with simple descriptors. Both MSD and bonding/antibonding states are two computationally efficient approaches for screening ultralow κ_L. We should mention that, given the high accuracy of harmonic dominant properties such as maximum MSD, we propose our model as a method to generate high quality data for direct prediction of phonon properties through other machine learning models, like prediction of extremely high or low lattice thermal conductivity materials. Physical insight into off-diagonal contribution to overall phonon transport is also analyzed with our model, demonstrating the general trend of high coherence contributions to the total thermal conductivity for low κ_L structures and the crossover from particle-like to wave-like heat conduction in diverse structures. The precise coherence contribution percentage depends on detailed phonon band structures and mode level phonon anharmonicity of different materials, rather than the single absolute value of κ_L. Our algorithm can grow and be adapted to even larger databases of unseen materials and is promising for accelerating discovery of novel crystals in emerging phonon mediated applications.

Acknowledgements

A.R. acknowledges the financial support by the Department of Energy, Office of Nuclear Energy, Integrated University Program Graduate Fellowship (IUP) under award no. DE-NE-0000095 and NASA SC Space Grant Consortium REAP Program (Award No.: 521383-RP-SC004). Research reported in this work was supported in part by NSF under awards 2030128 and 2110033, SC EPSCoR Program under award number (23-GC01), and an ASPIRE grant from the Office of the Vice President for Research at the University of South Carolina (project 80005046).

Author Contributions

M.H. conveyed the idea and designed and supervised the study. A.R. performed the neural network potential training and testing as well as the active learning loop. C.L. wrote the code utilizing compressive sensing lattice dynamics (CSLD) for 2nd and 3rd IFCs fitting. A.R. and C.S. performed DFT calculations. K.Y. revised the ShengBTE code to quantify the coherence contribution to overall thermal transport. M.A. performed the COHP calculations and bonding/antibonding analysis. A.R. prepared the draft of the manuscript. X.Z., H.Z., and M.H. revised the manuscript. All the authors contributed to discussions and interpretation of results in the manuscript.

Competing Interests

The authors declare no competing interests.

Chapter 6 Conclusion

In this work, machine learning-based atomic simulation was outlined for the development and demonstration of virtual high-throughput materials screening processes. This was first emphasized in Chapter 2 by the introduction of the SDNNFF model, whereby atomic forces were predicted for rapid DFT-level evaluations of displaced atoms in diamond-like materials for phonon dispersion and lattice thermal conductivity prediction. In short, the SDNNFF represents the local environment by subdividing the space surrounding atoms into grids. Density functions were centered at the grids and measure the influence of nearby atoms *via* decaying cosine functions, which collectively provided an image of the neighborhood to the neural network. As shown by the excellent agreement of the atomic forces and phonon properties, the atomic descriptors of the SDNNFF were sufficient for mimicking DFT force fields at ~10^3 times faster speeds.

In Chapter 3, molten salts LiF and FLiBe were modeled by existing DPSE in MD simulations for the prediction of thermophysical and dynamic properties, including density, thermal conductivity, and viscosity. Due to the difficulty of high temperature experiments of corrosive liquids, virtual experiment and measurement of molten salts was key for the development and understanding of heat transfer fluids for MSR design. As such, it was demonstrated that machine learning atomic potentials can capture the complex local structures in molten

salts as shown by the well-represented local structures and corresponding properties relative to experiment. However, this work emphasized the importance of verifying reference DFT calculations and models to represent the appropriate chemistries for the given system. In this case, the molten salts were shown to be highly sensitive to dispersion relations, which were not appropriately accommodated for in DFT *via* vdW corrections. This resulted in a 14% decrease of the melt density with respect to experiment for both LiF and FLiBe, in-turn displacing subsequent properties from their acceptable values. Insight into the training methodology and model construction for molten salt modeling was also provided, quantifying the required dataset sizes including various compressed and expanded volumes for stable molten salt potentials.

Because of the complex compositional space for molten salts, inspiration was drawn to develop general interatomic potentials to simultaneously represent a wide array of atomic systems independently of the exponential increase in model complexity. Additionally, training various atomic environments in a single model guarantees a level of transferability between similar chemistries, providing a generalized understanding of bonding behaviors. Henceforth, Chapter 4 continued the development of the SDNNFF with the Elemental-SDNNFF model by incorporation of weights corresponding to central and neighboring atoms. The performance of the Elemental-SDNNFF was demonstrated on 11,886 quaternary, half, and full Heusler structures including 55 different elements trained on a single network. Million-scale atomic environments were easily generated due to several factors, including 1) training on atomic forces providing *N* data per DFT

run, where *N* is the number of atoms in a simulation cell, 2) data augmentation of existing data providing a 3× boost to the dataset size, and 3) automatic proposal of new atomic environments from active learning iterations, which iteratively improved the model. As a result, excellent agreement was found in the phonon dispersions and lattice thermal conductivities with respect to DFT and experiments. Due to the several thousands of lattice thermal conductivity data, correlations were drawn to other physical properties, including volume and number density, and bonding/antibonding. Specifically, volume and number density scaled inversely and proportionally to the lattice thermal conductivity, respectively, whereas bonding/antibonding showed a competing influence. Interestingly, bonding/antibonding may be used to effectively filter out crystalline insulators, whereby many low lattice thermal conductivities were found at low bonding and high antibonding values. In this range, p-d orbital hybridization was observed whereby antibonding provides high phonon anharmonicity and low lattice thermal conductivity. Weyl points were also investigated whereby unique materials with topological phonons were screened by predicted IFCs.

To push the Elemental-SDNNFF further, Chapter 5 models 77,091 unexplored structures spanning 16 structural prototypes and 63 elements spanning the periodic table. With 27,059 structures predicted stable, lattice thermal conductivities were predicted with 13,461 structures owning < 1 $Wm^{-1}K^{-1}$. Other phonon-based properties were also predicted including the molar specific heat capacity, sound velocity, third order scattering channel volume, and thermal mean squared displacements. Because the thermal mean squared

displacements are computed with harmonic IFCs, they were shown as inexpensive descriptors for the lattice thermal conductivity. Physical insight into phonon hydrodynamics is also analyzed with the model, showing an inversely proportional trend of coherence phonons against total phonon thermal conductivity.

In conclusion, discovery of new materials for energy innovation is faster than ever before due to solidified chemistry and physics theory paired with advanced computing technologies. To satisfy the ever growing energy needs, new methodologies are necessary to accelerate understanding of nanoscale phenomena as related to micro- and macroscopic properties, as well as for rapid screening of materials for experimental synthesis and near-future deployment. Due to the abundance of data in the age of advanced computing, machine learning and artificial intelligence provides a pathway for representing complex phenomena otherwise difficult to solve or quantify through theory alone. As such, the methodologies presented here are a small yet critical part of a greater research effort to build upon rather than replace the current foundations of materials science, opening doors to a new level of understanding of materials and cutting-edge engineering.

References

[1] Eivari, H. A., Sohbatzadeh, Z., Mele, P., and Assadi, M. H. N., 2021, “Low Thermal Conductivity: Fundamentals and Theoretical Aspects in Thermoelectric Applications,” Mater. Today Energy, **21**, p. 100744.

[2] Ghosh, T., Dutta, M., Sarkar, D., and Biswas, K., 2022, “Insights into Low Thermal Conductivity in Inorganic Materials for Thermoelectrics,” J. Am. Chem. Soc., **2022**(1).

[3] Liu, Y., Zhao, T., Ju, W., Shi, S., Shi, S., and Shi, S., 2017, “Materials Discovery and Design Using Machine Learning,” J. Mater., **3**(3), pp. 159–177.

[4] Jones, R. O., 2015, “Density Functional Theory: Its Origins, Rise to Prominence, and Future,” Rev. Mod. Phys., **87**(3).

[5] Sendek, A. D., Cubuk, E. D., Antoniuk, E. R., Cheon, G., Cui, Y., and Reed, E. J., 2019, “Machine Learning-Assisted Discovery of Solid Li-Ion Conducting Materials,” Chem. Mater., **31**(2), pp. 342–352.

[6] Gheribi, A. E., Corradini, D., Dewan, L., Chartrand, P., Simon, C., Madden, P. A., and Salanne, M., 2014, “Prediction of the Thermophysical Properties of Molten Salt Fast Reactor Fuel from First-Principles,” Mol. Phys., **112**(9–10), pp. 1305–1312.

[7] Sun, C., Wen, B., and Bai, B., 2015, “Application of Nanoporous Graphene Membranes in Natural Gas Processing: Molecular Simulations of CH4/CO2, CH4/H2S and CH4/N2 Separation,” Chem. Eng. Sci., **138**, pp. 616–621.

[8] Qin, G., Qin, Z., Wang, H., and Hu, M., 2017, “Anomalously Temperature-Dependent Thermal Conductivity of Monolayer GaN with Large Deviations from the Traditional 1/T Law,” Phys. Rev. B, **95**(19), pp. 1–10.

[9] Salanne, M., Rotenberg, B., Jahn, S., Vuilleumier, R., Simon, C., and Madden, P. A., 2012, “Including Many-Body Effects in Models for Ionic Liquids,” Theor. Chem. Acc., **131**(3), pp. 1–16.

[10] Plimpton, S., 1997, “Short-Range Molecular Dynamics,” J. Comput. Phys., **117**(6), pp. 1–42.

[11] Cottenier, S., 2013, *Density Functional Theory and the Family of (L)APW-Methods: A Step-by-Step Introduction*.

[12] Friederich, P., Häse, F., Proppe, J., and Aspuru-Guzik, A., 2021, “Machine-Learned Potentials for next-Generation Matter Simulations,” Nat. Mater., **20**(6), pp. 750–761.

[13] Behler, J., 2015, “Constructing High-Dimensional Neural Network Potentials: A Tutorial Review,” Int. J. Quantum Chem., **115**(16), pp. 1032–1050.

[14] Behler, J., 2010, *Neural Network Potential-Energy Surfaces for Atomistic*

Simulations.

[15] Rodriguez, A., Lam, S., and Hu, M., 2021, “Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials,” ACS Appl. Mater. Interfaces.

[16] Behler, J., 2011, “Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials,” J. Chem. Phys., **134**(7).

[17] Artrith, N., and Behler, J., 2012, “High-Dimensional Neural Network Potentials for Metal Surfaces: A Prototype Study for Copper,” Phys. Rev. B - Condens. Matter Mater. Phys., **85**(4), pp. 1–13.

[18] Li, Z., Kermode, J. R., and De Vita, A., 2015, “Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces,” Phys. Rev. Lett., **114**(9), pp. 1–5.

[19] Smith, J. S., Isayev, O., and Roitberg, A. E., 2017, “ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost,” Chem. Sci., **8**(4), pp. 3192–3203.

[20] Li, W., Ando, Y., Minamitani, E., and Watanabe, S., 2017, “Study of Li Atom Diffusion in Amorphous Li3PO4 with Neural Network Potential,” J. Chem. Phys., **147**(21).

[21] Deringer, V. L., Bernstein, N., Bartók, A. P., Cliffe, M. J., Kerber, R. N., Marbella, L. E., Grey, C. P., Elliott, S. R., and Csányi, G., 2018, “Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven

Molecular Dynamics," J. Phys. Chem. Lett., **9**(11), pp. 2879–2885.

[22] Marques, M. R. G., Wolff, J., Steigemann, C., and Marques, M. A. L., 2019, "Neural Network Force Fields for Simple Metals and Semiconductors: Construction and Application to the Calculation of Phonons and Melting Temperatures," Phys. Chem. Chem. Phys., **21**(12), pp. 6506–6516.

[23] Minamitani, E., Ogura, M., and Watanabe, S., 2019, "Simulating Lattice Thermal Conductivity in Semiconducting Materials Using High-Dimensional Neural Network Potential," Appl. Phys. Express, **12**(9).

[24] Rowe, P., Csányi, G., Alfè, D., and Michaelides, A., 2018, "Development of a Machine Learning Potential for Graphene," Phys. Rev. B, **97**(5), pp. 1–12.

[25] Cramer, C. J., 2017, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons, Chichester.

[26] Xie, T., France-Lanord, A., Wang, Y., Shao-Horn, Y., and Grossman, J. C., 2019, "Graph Dynamical Networks for Unsupervised Learning of Atomic Scale Dynamics in Materials," Nat. Commun., **10**(1).

[27] Raff, L. M., Malshe, M., Hagan, M., Doughan, D. I., Rockley, M. G., and Komanduri, R., 2005, "Ab Initio Potential-Energy Surfaces for Complex, Multichannel Systems Using Modified Novelty Sampling and Feedforward Neural Networks," J. Chem. Phys., **122**(8), p. 084104.

[28] Jensen, J. H., 2010, *Molecular Modeling Basics*, CRC, Boca Raton, FL.

[29] Seminario, J. M., Ponce, V. H., Galvez-Aranda, D. E., Vicharra, C., Selis, L., Franco-Gallo, F., and Gamero T, M., 2018, “Molecular Dynamics Simulations of Solid Electrolytes for Li-Ion Nanobatteries,” ECS Meet. Abstr., **MA2018-02**(8), pp. 504–504.

[30] Schmidhuber, J., 2015, “Deep Learning in Neural Networks: An Overview,” Neural Networks, **61**, pp. 85–117.

[31] 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., 2013, “Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation,” APL Mater., **1**(1).

[32] Tanaka, I., Rajan, K., and Wolverton, C., 2018, “Data-Centric Science for Materials Innovation,” MRS Bull., **43**(9), pp. 659–663.

[33] Yang, H., Zhang, Z., Zhang, J., and Zeng, X. C., 2018, “Machine Learning and Artificial Neural Network Prediction of Interfacial Thermal Resistance between Graphene and Hexagonal Boron Nitride,” Nanoscale, **10**(40), pp. 19092–19099.

[34] Schmidt, J., Marques, M. R. G., Botti, S., and Marques, M. A. L., 2019, “Recent Advances and Applications of Machine Learning in Solid-State Materials Science,” npj Comput. Mater., **5**(1).

[35] Zhang, L., Han, J., Wang, H., Car, R., and Weinan, E., 2017, “Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics,” arXiv, pp. 1–7.

[36] Onat, B., Cubuk, E. D., Malone, B. D., and Kaxiras, E., 2018, “Implanted Neural Network Potentials: Application to Li-Si Alloys,” Phys. Rev. B, **97**(9), pp. 1–9.

[37] Behler, J., and Parrinello, M., 2007, “Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces,” Phys. Rev. Lett., **98**(14), pp. 1–4.

[38] Singraber, A., Behler, J., and Dellago, C., 2019, “Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials,” J. Chem. Theory Comput., **15**(3), pp. 1827–1840.

[39] Parlinski, K., Li, Z. Q., and Kawazoe, Y., 1997, “First-Principles Determination of the Soft Mode in Cubic ZrO2,” Phys. Rev. Lett., **78**(21), pp. 4063–4066.

[40] Tadano, T., and Tsuneyuki, S., 2018, “First-Principles Lattice Dynamics Method for Strongly Anharmonic Crystals,” J. Phys. Soc. Japan, **87**(4), pp. 1–11.

[41] Huang, Y., Kang, J., Goddard, W. A., and Wang, L. W., 2019, “Density Functional Theory Based Neural Network Force Fields from Energy Decompositions,” Phys. Rev. B, **99**(6), pp. 1–11.

[42] Van Der Walt, S., Colbert, S. C., and Varoquaux, G., 2011, “The NumPy Array: A Structure for Efficient Numerical Computation,” Comput. Sci. Eng., **13**(2), pp. 22–30.

[43] Abadi, M., Isard, M., and Murray, D. G., 2017, “A Computational Model for TensorFlow an Introduction,” MAPL 2017 - Proc. 1st ACM SIGPLAN Int. Work. Mach. Learn. Program. Lang. co-located with PLDI 2017, pp. 1–7.

[44] He, K., Zhang, X., Ren, S., and Sun, J., 2015, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, pp. 1026–1034.

[45] Kingma, D. P., and Ba, J. L., 2015, “Adam: A Method for Stochastic Optimization,” 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–15.

[46] Clevert, D. A., Unterthiner, T., and Hochreiter, S., 2016, “Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs),” 4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc., pp. 1–14.

[47] Togo, A., and Tanaka, I., 2015, “First Principles Phonon Calculations in Materials Science,” Scr. Mater., **108**, pp. 1–5.

[48] Togo, A., Chaput, L., and Tanaka, I., 2015, “Distributions of Phonon Lifetimes in Brillouin Zones,” Phys. Rev. B - Condens. Matter Mater. Phys., **91**(9).

[49] McGaughey, A. J. H., Jain, A., Kim, H. Y., and Fu, B., 2019, “Phonon Properties and Thermal Conductivity from First Principles, Lattice Dynamics, and the Boltzmann Transport Equation,” J. Appl. Phys., **125**(1).

[50] Dong, J., and Chen, A.-B., 2004, “Fundamental Properties of SiC: Crystal Structure, Bonding Energy, Band Structure, and Lattice Vibrations,” pp. 63–87.

[51] Ho, C. Y., Powell, R. W., and Liley, P. E., 1972, “Thermal Conductivity of the Elements,” J. Phys. Chem. Ref. Data, **1**(2), pp. 279–421.

[52] Kulda, J., Strauch, D., Pavone, P., and Ishii, Y., 1994, “Inelastic-Neutron-Scattering Study of Phonon Eigenvectors and Frequencies in Si,” Phys. Rev. B, **50**(18), pp. 13347–13354.

[53] Onn, D. G., Witek, A., Qiu, Y. Z., Anthony, T. R., and Banholzer, W. F., 1992, “Some Aspects of the Thermal Conductivity of Isotopically Enriched Diamond Single Crystals,” Phys. Rev. Lett., **68**(18), pp. 2806–2809.

[54] Warren, J. L., Yarnell, J. L., Dolling, G., and Cowley, R. A., 1967, “Lattice Dynamics of Diamond,” Phys. Rev., **158**(3), pp. 805–808.

[55] Schwoerer-Böhning, M., Macrander, A. T., and Arms, D. A., 1998, “Phonon Dispersion of Diamond Measured by Inelastic X-Ray Scattering,” Phys. Rev. Lett., **80**(25), pp. 5572–5575.

[56] Kulda, J., Kainzmaier, H., Strauch, D., Dorner, B., Lorenzen, M., and Krisch, M., 2002, “Overbending of the Longitudinal Optical Phonon Branch in Diamond as Evidenced by Inelastic Neutron and X-Ray Scattering,” Phys. Rev. B - Condens. Matter Mater. Phys., **66**(24), pp. 1–4.

[57] Sparavigna, A., 2002, “Lattice Thermal Conductivity in Cubic Silicon

Carbide," Phys. Rev. B - Condens. Matter Mater. Phys., **66**(17), pp. 1–5.

[58] Serrano, J., Strempfer, J., Cardona, M., Schwoerer-Böhning, M., Requardt, H., Lorenzen, M., Stojetz, B., Pavone, P., and Choyke, W. J., 2002, "Determination of the Phonon Dispersion of Zinc Blende (3C) Silicon Carbide by Inelastic x-Ray Scattering," Appl. Phys. Lett., **80**(23), pp. 4360–4362.

[59] Widulle, F., Ruf, T., Buresch, O., Debernardi, A., and Cardona, M., 1999, "Raman Study of Isotope Effects and Phonon Eigenvectors in SiC," Phys. Rev. Lett., **82**(15), pp. 3089–3092.

[60] Ma, H., Li, C., Tang, S., Yan, J., Alatas, A., Lindsay, L., Sales, B. C., and Tian, Z., 2016, "Boron Arsenide Phonon Dispersion from Inelastic X-Ray Scattering: Potential for Ultrahigh Thermal Conductivity," Phys. Rev. B, **94**(22), pp. 1–4.

[61] Greene, R. G., Luo, H., Ruoff, A. L., Trail, S. S., and DiSalvo, F. J., 1994, "Pressure Induced Metastable Amorphization of BAs: Evidence for a Kinetically Frustrated Phase Transformation," Phys. Rev. Lett., **73**(18), pp. 2476–2479.

[62] Kang, J. S., Li, M., Wu, H., Nguyen, H., and Hu, Y., 2018, "Experimental Observation of High Thermal Conductivity in Boron Arsenide," Science (80-.)., **361**(6402), pp. 575–578.

[63] Maaten, L. van der, and Hinton, G., 2008, "Visualizing Data Using T-SNE," J. Mach. Learn. Res., **9**(2579–2605), p. 85.

[64] Müller-Plathe, F., 1997, “A Simple Nonequilibrium Molecular Dynamics Method for Calculating the Thermal Conductivity,” J. Chem. Phys., **106**(14), pp. 6082–6085.

[65] Wang, H., Ji, X., Chen, C., Xu, K., and Miao, L., 2013, “Lithium Diffusion in Silicon and Induced Structure Disorder: A Molecular Dynamics Study,” AIP Adv., **3**(11).

[66] Gastegger, M., Schwiedrzik, L., Bittermann, M., Berzsenyi, F., and Marquetand, P., 2018, “WACSF - Weighted Atom-Centered Symmetry Functions as Descriptors in Machine Learning Potentials,” J. Chem. Phys., **148**(24).

[67] Bochkarev, A. S., Van Roekeghem, A., Mossa, S., and Mingo, N., 2019, “Anharmonic Thermodynamics of Vacancies Using a Neural Network Potential,” Phys. Rev. Mater., **3**(9), pp. 1–7.

[68] 2014, “A Technology Roadmap for Generation IV Nuclear Energy Systems,” Nucl. Energy Res. Advis. Comm. Gener. IV Int. Forum, p. GIF--002--00.

[69] Andrew, C. K., 2017, “A Comparison of Advanced Nuclear Technologies,” Columbia SIPA, p. 112.

[70] Heuer, D., Merle-Lucotte, E., Allibert, M., Brovchenko, M., Ghetta, V., and Rubiolo, P., 2014, “Towards the Thorium Fuel Cycle with Molten Salt Fast Reactors,” Ann. Nucl. Energy, **64**, pp. 421–429.

[71] Elsheikh, B. M., 2013, “Safety Assessment of Molten Salt Reactors in Comparison with Light Water Reactors,” J. Radiat. Res. Appl. Sci., **6**(2), pp. 63–70.

[72] Carlson, K., Gardner, L., Moon, J., Riley, B., Amoroso, J., and Chidambaram, D., 2021, “Molten Salt Reactors and Electrochemical Reprocessing: Synthesis and Chemical Durability of Potential Waste Forms for Metal and Salt Waste Streams,” Int. Mater. Rev., **66**(5), pp. 339–363.

[73] Yoshioka, R., Kinoshita, M., and Scott, I., 2017, *Materials*.

[74] 2017, “Molten Salt Chemistry Workshop. Innovation Approaches to Accelerate Molten Salt Reactor Development and Deployment,” *Oak Ridge National Laboratory. Technology and Applied R&D*.

[75] Dai, J. X., Zhang, W., Ren, C. L., and Guo, X. J., 2020, “Prediction of Dynamics Properties of ThF4-Based Fluoride Molten Salts by Molecular Dynamic Simulation,” J. Mol. Liq., **318**, p. 114059.

[76] Cadwallader, L. C., and Longhurst, G. R., 1999, “Flibe Use in Fusion Reactors: An Initial Safety Assessment,” Idaho Natl. Eng. Environ. Lab., (March), pp. 1–25.

[77] Dewan, L. C., Simon, C., Madden, P. A., Hobbs, L. W., and Salanne, M., 2013, “Molecular Dynamics Simulation of the Thermodynamic and Transport Properties of the Molten Salt Fast Reactor Fuel LiF-ThF4,” J. Nucl. Mater., **434**(1–3), pp. 322–327.

[78] Salanne, M., Simon, C., Turq, P., Heaton, R. J., and Madden, P. A., 2006, “A First-Principles Description of Liquid BeF2 and Its Mixtures with LiF: 2. Network Formation in LiF-BeF2,” J. Phys. Chem. B, **110**(23), pp. 11461–11467.

[79] Sarou-Kanian, V., Rollet, A. L., Salanne, M., Simon, C., Bessada, C., and Madden, P. A., 2009, “Diffusion Coefficients and Local Structure in Basic Molten Fluorides: In Situ NMR Measurements and Molecular Dynamics Simulations,” Phys. Chem. Chem. Phys., **11**(48), pp. 11501–11506.

[80] Ishii, Y., Kasai, S., Salanne, M., and Ohtori, N., 2015, “Transport Coefficients and the Stokes-Einstein Relation in Molten Alkali Halides with Polarisable Ion Model,” Mol. Phys., **113**(17–18), pp. 2442–2450.

[81] Salanne, M., and Madden, P. A., 2011, “Polarization Effects in Ionic Solids and Melts,” Mol. Phys., **109**(19), pp. 2299–2315.

[82] Ishii, Y., Sato, K., Salanne, M., Madden, P. A., and Ohtori, N., 2014, “Thermal Conductivity of Molten Alkali Metal Fluorides (LiF, NaF, KF) and Their Mixtures,” J. Phys. Chem. B, **118**(12), pp. 3385–3391.

[83] Nam, H. O., Bengtson, A., Vörtler, K., Saha, S., Sakidja, R., and Morgan, D., 2014, “First-Principles Molecular Dynamics Modeling of the Molten Fluoride Salt with Cr Solute,” J. Nucl. Mater., **449**(1–3), pp. 148–157.

[84] Lam, S. T., Li, Q. J., Mailoa, J., Forsberg, C., Ballinger, R., and Li, J., 2021, “The Impact of Hydrogen Valence on Its Bonding and Transport in Molten Fluoride Salts,” J. Mater. Chem. A, **9**(3), pp. 1784–1794.

[85] Lam, S. T., Li, Q.-J., Ballinger, R., Forsberg, C., and Li, J., 2021, “Modeling LiF and FLiBe Molten Salts with Robust Neural Network Interatomic Potential,” ACS Appl. Mater. Interfaces.

[86] Zhang, L., Han, J., Wang, H., Car, R., and Weinan, E., 2018, “Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics,” Phys. Rev. Lett., **120**(14), pp. 1–6.

[87] Jabes, B. S., Agarwal, M., and Chakravarty, C., 2012, “Structure and Transport Properties of LiF-BeF2 Mixtures: Comparison of Rigid and Polarizable Ion Potentials,” J. Chem. Sci., **124**(1), pp. 261–269.

[88] Gheribi, A. E., Torres, J. A., and Chartrand, P., 2014, “Recommended Values for the Thermal Conductivity of Molten Salts between the Melting and Boiling Points,” Sol. Energy Mater. Sol. Cells, **126**, pp. 11–25.

[89] Khokhlov, V., Ignatiev, V., and Afonichkin, V., 2009, “Evaluating Physical Properties of Molten Salt Reactor Fluoride Mixtures,” J. Fluor. Chem., **130**(1), pp. 30–37.

[90] Wang, J., Sun, Z., Lu, G., and Yu, J., 2014, “Molecular Dynamics Simulations of the Local Structures and Transport Coefficients of Molten Alkali Chlorides,” J. Phys. Chem. B, **118**(34), pp. 10196–10206.

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

[92] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., 2016, “TensorFlow: A System for Large-Scale Machine Learning,” Proc. 12th USENIX Symp. Oper. Syst. Des. Implementation, OSDI 2016, pp. 265–283.

[93] Zhang, L., Han, J., Wang, H., Saidi, W. A., Car, R., and Weinan, E., 2018, “End-to-End Symmetry Preserving Inter-Atomic Potential Energy Model for Finite and Extended Systems,” Adv. Neural Inf. Process. Syst., **2018-Decem**(NeurIPS 2018), pp. 4436–4446.

[94] Adams, D. J., and McDonald, I. R., 1974, “Solid State Physics Related Content Rigid-Ion Models of the Interionic Potential in the Alkali Halides,” J. Phys. C Solid State Phys., **7**(16), pp. 2761–2775.

[95] Wang, S., Luo, H., Deng, H., Xiao, S., and Hu, W., 2017, “A Molecular Dynamics Study of the Transport Properties of LiF-BeF2-ThF4 Molten Salt,” J. Mol. Liq., **234**, pp. 220–226.

[96] Rodriguez, A., Liu, Y., and Hu, M., 2020, “Spatial Density Neural Network Force Fields with First-Principles Level Accuracy and Application to Thermal Transport,” Phys. Rev. B, **102**(3), p. 35203.

[97] Ohtori, N., Oono, T., and Takase, K., 2009, “Thermal Conductivity of Molten Alkali Halides: Temperature and Density Dependence,” J. Chem. Phys., **130**(4).

[98] Janz, G., 1988, “Thermodynamic and Transport Properties for Molten Salts Correlation Equations for Critically Evaluated Density, Surface Tension, Electrical Conductance, and Viscosity Data,” J. Phys. Chem. Ref. Data, **17**(2), pp. 1–309.

[99] Smith, A. L., Capelli, E., Konings, R. J. M., and Gheribi, A. E., 2020, “A New Approach for Coupled Modelling of the Structural and Thermo-Physical Properties of Molten Salts. Case of a Polymeric Liquid LiF-BeF2,” J. Mol. Liq., **299**, p. 112165.

[100] Ko, T. W., Finkler, J. A., Goedecker, S., and Behler, J., 2021, “A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-Local Charge Transfer,” Nat. Commun., **12**(1), pp. 1–11.

[101] Luo, H., Xiao, S., Wang, S., Huai, P., Deng, H., and Hu, W., 2016, “Molecular Dynamics Simulation of Diffusion and Viscosity of Liquid Lithium Fluoride,” Comput. Mater. Sci., **111**, pp. 203–208.

[102] Galamba, N., De Castro, C. A. N., and Ely, J. F., 2004, “Thermal Conductivity of Molten Alkali Halides from Equilibrium Molecular Dynamics Simulations,” J. Chem. Phys., **120**(18), pp. 8676–8682.

[103] Galamba, N., De Castro, C. A. N., and Ely, J. F., 2007, “Equilibrium and Nonequilibrium Molecular Dynamics Simulations of the Thermal Conductivity of Molten Alkali Halides,” J. Chem. Phys., **126**(20).

[104] Urquiza, M. L., Islam, M. M., Van Duin, A. C. T., Cartoixà, X., and

Strachan, A., 2021, “Atomistic Insights on the Full Operation Cycle of a HfO2-Based Resistive Random Access Memory Cell from Molecular Dynamics,” ACS Nano, **15**(8), pp. 12945–12954.

[105] Golyshev. V.D. ; Gonik, M. A., 1992, “High-Temperature Thermophysical Properties of Non-Scattering Semitransparent Materials III: Thermal Conductivity of Melts.,” High Temp. - High Press., **24**(6), pp. 677–688.

[106] Golyshev, V.D.; Gonik, M.A.; Petrov, V.A.; Putilin, Y. M., 1983, “Teplofiz,” Temp. 21, pp. 899–903.

[107] Gierszewski, P., Mikic, B., and Trodreas, N., 1980, “Properties Correlations for Li Na He FLIBE and Water in Fusion Reactor Applications,” Massachusetts Inst. Technol.

[108] Kato, Y.; Furukawa, K.; Araki, N.; Kobayasi, K., 1983, “Thermal Diffusivity Measurement of Molten Salts by Use of a Simple Ceramic Cell,” High Temp. - High Press., pp. 191–198.

[109] Zhang, Y., Otani, A., and Maginn, E. J., 2015, “Reliable Viscosity Calculation from Equilibrium Molecular Dynamics Simulations: A Time Decomposition Method,” J. Chem. Theory Comput., **11**(8), pp. 3537–3546.

[110] Janz, G. J., Allen, C. B., Downey, Jr., J. R., and Tomkins, R. P. T., 1978, “Physical Properties Data Compilations Relevant to Energy Storage.”

[111] Popescu, A. M., and Constantin, V., 2015, “Viscosity of Alkali Fluoride Ionic Melts at Temperatures up to 373.15 K above Melting Points,” Chem. Eng.

Commun., **202**(12), pp. 1703–1710.

[112] Abe, Y., Kosugiyama, O., and Nagashima, A., 1981, “Viscosity of LiF-BeF2 Eutectic Mixture (XBeF2 = 0.328) and LiF Single Salt at Elevated Temperatures,” J. Nucl. Mater., **99**(2–3), pp. 173–183.

[113] Zhang, L., Han, J., Wang, H., Saidi, W. A., Car, R., and Weinan, E., 2018, “End-to-End Symmetry Preserving Inter-Atomic Potential Energy Model for Finite and Extended Systems,” Adv. Neural Inf. Process. Syst., **2018-Decem**, pp. 4436–4446.

[114] Lu, D., Wang, H., Chen, M., Lin, L., Car, R., E, W., Jia, W., and Zhang, L., 2021, “86 PFLOPS Deep Potential Molecular Dynamics Simulation of 100 Million Atoms with Ab Initio Accuracy,” Comput. Phys. Commun., **259**.

[115] Zhang, L., Lin, D. Y., Wang, H., Car, R., and Weinan, E., 2019, “Active Learning of Uniformly Accurate Interatomic Potentials for Materials Simulation,” Phys. Rev. Mater., **3**(2), p. 23804.

[116] Xia, Y., Hegde, V. I., Pal, K., Hua, X., Gaines, D., Patel, S., He, J., Aykol, M., and Wolverton, C., 2020, “High-Throughput Study of Lattice Thermal Conductivity in Binary Rocksalt and Zinc Blende Compounds Including Higher-Order Anharmonicity,” Phys. Rev. X, **10**(4), p. 41029.

[117] Roekeghem, A., Carrete, J., Oses, C., Curtarolo, S., and Mingo, N., 2016, “High-Throughput Computation of Thermal Conductivity of High-Temperature Solid Phases: The Case of Oxide and Fluoride Perovskites,” Phys. Rev. X, **6**(4), pp. 1–10.

[118] Ju, S., Yoshida, R., Liu, C., Wu, S., Hongo, K., Tadano, T., and Shiomi, J., 2021, “Exploring Diamondlike Lattice Thermal Conductivity Crystals via Feature-Based Transfer Learning,” Phys. Rev. Mater., **5**(5), pp. 1–2.

[119] Yamada, H., Liu, C., Wu, S., Koyama, Y., Ju, S., Shiomi, J., Morikawa, J., and Yoshida, R., 2019, “Predicting Materials Properties with Little Data Using Shotgun Transfer Learning,” ACS Cent. Sci., **5**(10), pp. 1717–1730.

[120] Miyazaki, H., Tamura, T., Mikami, M., Watanabe, K., Ide, N., Ozkendir, O. M., and Nishino, Y., 2021, “Machine Learning Based Prediction of Lattice Thermal Conductivity for Half-Heusler Compounds Using Atomic Information,” Sci. Rep., **11**(1), pp. 1–8.

[121] Zhu, T., He, R., Gong, S., Xie, T., Gorai, P., Nielsch, K., and Grossman, J. C., 2021, “Charting Lattice Thermal Conductivity for Inorganic Crystals and Discovering Rare Earth Chalcogenides for Thermoelectrics,” Energy Environ. Sci., **14**(6), pp. 3559–3566.

[122] Schleder, G. R., Padilha, A. C. M., Acosta, C. M., Costa, M., and Fazzio, A., 2019, “From DFT to Machine Learning: Recent Approaches to Materials Science - A Review,” JPhys Mater., **2**(3).

[123] Simoncelli, M., Marzari, N., and Mauri, F., 2019, “Unified Theory of Thermal Transport in Crystals and Glasses,” Nat. Phys., **15**(8), pp. 809–813.

[124] Han, Z., Yang, X., Li, W., Feng, T., and Ruan, X., 2022, “FourPhonon: An Extension Module to ShengBTE for Computing Four-Phonon Scattering

Rates and Thermal Conductivity," Comput. Phys. Commun., **270**, p. 108179.

[125] Bartók, A. P., Kermode, J., Bernstein, N., and Csányi, G., 2018, "Machine Learning a General-Purpose Interatomic Potential for Silicon," Phys. Rev. X, **8**(4), p. 41048.

[126] Bartók, A. P., Payne, M. C., Kondor, R., and Csányi, G., 2010, "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons," Phys. Rev. Lett., **104**(13), pp. 1–4.

[127] Xie, H., Gu, X., and Bao, H., 2017, "Effect of the Accuracy of Interatomic Force Constants on the Prediction of Lattice Thermal Conductivity," Comput. Mater. Sci., **138**, pp. 368–376.

[128] Carrete, J., Li, W., Mingo, N., Wang, S., and Curtarolo, S., 2014, "Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling," Phys. Rev. X, **4**(1), pp. 1–9.

[129] Seko, A., Togo, A., Hayashi, H., Tsuda, K., Chaput, L., and Tanaka, I., 2015, "Prediction of Low-Thermal-Conductivity Compounds with First-Principles Anharmonic Lattice-Dynamics Calculations and Bayesian Optimization," Phys. Rev. Lett., **115**(20), pp. 1–5.

[130] Armitage, N. P., Mele, E. J., and Vishwanath, A., 2018, "Weyl and Dirac Semimetals in Three-Dimensional Solids," Rev. Mod. Phys., **90**(1), p. 015001.

[131] Zhang, T., Song, Z., Alexandradinata, A., Weng, H., Fang, C., Lu, L., and Fang, Z., 2018, “Double-Weyl Phonons in Transition-Metal Monosilicides,” Phys. Rev. Lett., **120**(1), p. 016401.

[132] Tang, D. S., and Cao, B. Y., 2021, “Topological Effects of Phonons in GaN and AlGaN: A Potential Perspective for Tuning Phonon Transport,” J. Appl. Phys., **129**(8), p. 085102.

[133] Wan, X., Turner, A. M., Vishwanath, A., and Savrasov, S. Y., 2011, “Topological Semimetal and Fermi-Arc Surface States in the Electronic Structure of Pyrochlore Iridates,” Phys. Rev. B - Condens. Matter Mater. Phys., **83**(20), p. 205101.

[134] Saal, J. E., Kirklin, S., Aykol, M., Meredig, B., and Wolverton, C., 2013, “Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD),” Jom, **65**(11), pp. 1501–1509.

[135] Kresse, G., and Furthmüller, J., 1996, “Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set,” Comput. Mater. Sci., **6**(1), pp. 15–50.

[136] Kresse, G., and Hafner, J., 1993, “Ab Initio Molecular Dynamics for Liquid Metals,” Phys. Rev. B, **47**(1), pp. 558–561.

[137] Kresse, G., and Furthmüller, J., 1996, “Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set,” Phys. Rev. B, **54**(16), pp. 11169–11186.

[138] Zhou, F., Nielson, W., Xia, Y., and Ozoliņš, V., 2019, “Compressive Sensing Lattice Dynamics. I. General Formalism,” Phys. Rev. B, **100**(18).

[139] Li, W., Carrete, J., Katcho, N. A., and Mingo, N., 2014, “ShengBTE: A Solver of the Boltzmann Transport Equation for Phonons,” Comput. Phys. Commun., **185**(6), pp. 1747–1758.

[140] Soluyanov, A. A., and Vanderbilt, D., 2011, “Computing Topological Invariants without Inversion Symmetry,” Phys. Rev. B - Condens. Matter Mater. Phys., **83**(23), p. 235401.

[141] Wu, Q. S., Zhang, S. N., Song, H. F., Troyer, M., and Soluyanov, A. A., 2018, “WannierTools: An Open-Source Software Package for Novel Topological Materials,” Comput. Phys. Commun., **224**, pp. 405–416.

[142] Gonze, X., and Lee, C., 1997, “Dynamical Matrices, Born Effective Charges, Dielectric Permittivity Tensors, and Interatomic Force Constants from Density-Functional Perturbation Theory,” Phys. Rev. B - Condens. Matter Mater. Phys., **55**(16), pp. 10355–10368.

[143] Shen, Q., Chen, L., Goto, T., Hirai, T., Yang, J., Meisner, G. P., and Uher, C., 2001, “Effects of Partial Substitution of Ni by Pd on the Thermoelectric Properties of ZrNiSn-Based Half-Heusler Compounds,” Appl. Phys. Lett., **79**(25), pp. 4165–4167.

[144] Zou, D. F., Xie, S. H., Liu, Y. Y., Lin, J. G., and Li, J. Y., 2013, “Electronic Structure and Thermoelectric Properties of Half-Heusler Zr 0.5Hf0.5NiSn by First-Principles Calculations,” J. Appl. Phys., **113**(19).

[145] Hermet, P., and Jund, P., 2016, “Lattice Thermal Conductivity of NiTiSn Half-Heusler Thermoelectric Materials from First-Principles Calculations,” J. Alloys Compd., **688**, pp. 248–252.

[146] He, R., Zhu, T., Wang, Y., Wolff, U., Jaud, J. C., Sotnikov, A., Potapov, P., Wolf, D., Ying, P., Wood, M., Liu, Z., Feng, L., Rodriguez, N. P., Snyder, G. J., Grossman, J. C., Nielsch, K., and Schierning, G., 2020, “Unveiling the Phonon Scattering Mechanisms in Half-Heusler Thermoelectric Compounds,” Energy Environ. Sci., **13**(12), pp. 5165–5176.

[147] Wang, S., Wang, Z., Setyawan, W., Mingo, N., and Curtarolo, S., 2011, “Assessing the Thermoelectric Properties of Sintered Compounds via High-Throughput Ab-Initio Calculations,” Phys. Rev. X, **1**(2), pp. 1–8.

[148] Xia, Y., Bhattacharya, S., Ponnambalam, V., Pope, A. L., Poon, S. J., and Tritt, T. M., 2000, “Thermoelectric Properties of Semimetallic (Zr, Hf)CoSb Half-Heusler Phases,” J. Appl. Phys., **88**(4), pp. 1952–1955.

[149] Zhou, M., Chen, L., Zhang, W., and Feng, C., 2005, “Disorder Scattering Effect on the High-Temperature Lattice Thermal Conductivity of TiCoSb-Based Half-Heusler Compounds,” J. Appl. Phys., **98**(1).

[150] Yu, C., Zhu, T. J., Shi, R. Z., Zhang, Y., Zhao, X. B., and He, J., 2009, “High-Performance Half-Heusler Thermoelectric Materials Hf1-x ZrxNiSn1-YSby Prepared by Levitation Melting and Spark Plasma Sintering,” Acta Mater., **57**(9), pp. 2757–2764.

[151] Zhu, H., Mao, J., Li, Y., Sun, J., Wang, Y., Zhu, Q., Li, G., Song, Q., Zhou,

J., Fu, Y., He, R., Tong, T., Liu, Z., Ren, W., You, L., Wang, Z., Luo, J., Sotnikov, A., Bao, J., Nielsch, K., Chen, G., Singh, D. J., and Ren, Z., 2019, “Discovery of TaFeSb-Based Half-Heuslers with High Thermoelectric Performance,” Nat. Commun., **10**(1), pp. 1–8.

[152] Bhattacharya, S., Skove, M. J., Russell, M., Tritt, T. M., Xia, Y., Ponnambalam, V., Poon, S. J., and Thadhani, N., 2008, “Effect of Boundary Scattering on the Thermal Conductivity of TiNiSn-Based Half-Heusler Alloys,” Phys. Rev. B - Condens. Matter Mater. Phys., **77**(18), pp. 1–8.

[153] Shastri, S. S., and Pandey, S. K., “First-Principles Electronic Structure , Phonon Properties , Lattice Thermal Conductivity and Prediction of Figure of Merit of FeVSb Half-Heusler.”

[154] Lue, C. S., Chen, C. F., Lin, J. Y., Yu, Y. T., and Kuo, Y. K., 2007, “Thermoelectric Properties of Quaternary Heusler Alloys Fe2 VAl1-x Six,” Phys. Rev. B - Condens. Matter Mater. Phys., **75**(6), pp. 1–6.

[155] Yan, J., Liu, F., Ma, G., Gong, B., Zhu, J., Wang, X., Ao, W., Zhang, C., Li, Y., and Li, J., 2018, “Suppression of the Lattice Thermal Conductivity in NbFeSb-Based Half-Heusler Thermoelectric Materials through High Entropy Effects,” Scr. Mater., **157**, pp. 129–134.

[156] Jollife, I. T., and Cadima, J., 2016, “Principal Component Analysis: A Review and Recent Developments,” Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., **374**(2065).

[157] Slack, G. A., 1973, “Nonmetallic Crystals with High Thermal Conductivity,” J. Phys. Chem. Solids, **34**(2), pp. 321–335.

[158] Keyes, R. W., 1959, “High-Temperature Thermal Conductivity of Insulating Crystals: Relationship to the Melting Point,” Phys. Rev., **115**(3), pp. 564–567.

[159] He, J., Xia, Y., Lin, W., Pal, K., Zhu, Y., Kanatzidis, M. G., and Wolverton, C., 2021, “Accelerated Discovery and Design of Ultralow Lattice Thermal Conductivity Materials Using Chemical Bonding Principles,” Adv. Funct. Mater., **2108532**, pp. 1–15.

[160] Jaffe, J. E., and Zunger, A., 1984, “Theory of the Band-Gap Anomaly in ABC2 Chalcopyrite Semiconductors,” Phys. Rev. B, **29**(4), pp. 1882–1906.

[161] Wei, S. H., and Zunger, A., 1988, “Role of Metal d States in II-VI Semiconductors,” Phys. Rev. B, **37**(15), pp. 8958–8981.

[162] Zhang, T., Jiang, Y., Song, Z., Huang, H., He, Y., Fang, Z., Weng, H., and Fang, C., 2019, “Catalogue of Topological Electronic Materials,” Nat. 2019 5667745, **566**(7745), pp. 475–479.

[163] Li, J., Liu, J., Baronett, S. A., Liu, M., Wang, L., Li, R., Chen, Y., Li, D., Zhu, Q., and Chen, X. Q., 2021, “Computation and Data Driven Discovery of Topological Phononic Materials,” Nat. Commun. 2021 121, **12**(1), pp. 1–12.

[164] Huang, S. M., Xuc, S. Y., Belopolski, I., Lee, C. C., Chang, G., Chang, T. R., Wang, B. K., Alidoust, N., Bian, G., Neupane, M., Sanchez, D., Zheng,

H., Jeng, H. T., Bansil, A., Neupert, T., Lin, H., and Hasan, M. Z., 2016, “New Type of Weyl Semimetal with Quadratic Double Weyl Fermions,” Proc. Natl. Acad. Sci. U. S. A., **113**(5), pp. 1180–1185.

[165] Padture, N. P., Gell, M., and Jordan, E. H., 2002, “Thermal Barrier Coatings for Gas-Turbine Engine Applications,” Science (80-.)., **296**(5566), pp. 280–284.

[166] Dumur, Satzinger, K. J., Peairs, G. A., Chou, M. H., Bienfait, A., Chang, H. S., Conner, C. R., Grebel, J., Povey, R. G., Zhong, Y. P., and Cleland, A. N., 2021, “Quantum Communication with Itinerant Surface Acoustic Wave Phonons,” npj Quantum Inf., **7**(1), pp. 1–5.

[167] Tang, D. S., Qin, G. Z., Hu, M., and Cao, B. Y., 2020, “Thermal Transport Properties of GaN with Biaxial Strain and Electron-Phonon Coupling,” J. Appl. Phys., **127**(3).

[168] Qin, G., and Hu, M., 2018, “Accelerating Evaluation of Converged Lattice Thermal Conductivity,” npj Comput. Mater., **4**(1).

[169] Zhou, Y., Xiong, S., Zhang, X., Volz, S., and Hu, M., 2018, “Thermal Transport Crossover from Crystalline to Partial-Crystalline Partial-Liquid State,” Nat. Commun., **9**(1), pp. 1–8.

[170] Qin, G., Qin, Z., Wang, H., and Hu, M., 2018, “Lone-Pair Electrons Induced Anomalous Enhancement of Thermal Transport in Strained Planar Two-Dimensional Materials,” Nano Energy, **50**(May), pp. 425–430.

[171] Lindsay, L., Katre, A., Cepellotti, A., and Mingo, N., 2019, “Perspective on Ab Initio Phonon Thermal Transport,” J. Appl. Phys., **126**(5), pp. 1–20.

[172] Li, J. F., Liu, W. S., Zhao, L. D., and Zhou, M., 2010, “High-Performance Nanostructured Thermoelectric Materials,” NPG Asia Mater., **2**(4), pp. 152–158.

[173] Ma, T., Chakraborty, P., Guo, X., Cao, L., and Wang, Y., 2020, *First-Principles Modeling of Thermal Transport in Materials: Achievements, Opportunities, and Challenges*, Springer US.

[174] Anand, S., Wood, M., Xia, Y., Wolverton, C., and Snyder, G. J., 2019, “Double Half-Heuslers,” Joule, **3**(5), pp. 1226–1238.

[175] He, J., Amsler, M., Xia, Y., Naghavi, S. S., Hegde, V. I., Hao, S., Goedecker, S., Ozoliņš, V., and Wolverton, C., 2016, “Ultralow Thermal Conductivity in Full Heusler Semiconductors,” Phys. Rev. Lett., **117**(4), pp. 1–6.

[176] Ding, J., Lanigan-Atkins, T., Calderón-Cueva, M., Banerjee, A., Abernathy, D. L., Said, A., Zevalkink, A., and Delaire, O., 2021, “Soft Anharmonic Phonons and Ultralow Thermal Conductivity in Mg3(Sb, Bi)2 Thermoelectrics,” Sci. Adv., **7**(21), pp. 1–8.

[177] Pal, K., Xia, Y., He, J., and Wolverton, C., 2019, “Intrinsically Low Lattice Thermal Conductivity Derived from Rattler Cations in an AMM′Q 3 Family of Chalcogenides,” Chem. Mater., **31**(21), pp. 8734–8741.

[178] Pal, K., Xia, Y., Shen, J., He, J., Luo, Y., Kanatzidis, M. G., and Wolverton, C., 2021, “Accelerated Discovery of a Large Family of Quaternary Chalcogenides with Very Low Lattice Thermal Conductivity,” npj Comput. Mater., **7**(1).

[179] Zhu, Z., Xi, J., and Yang, J., 2022, “Significant Reduction in Lattice Thermal Conductivity in a P-Type Filled Skutterudite Due to Strong Electron-Phonon Interactions,” J. Mater. Chem. A, **10**(25), pp. 13484–13491.

[180] Li, C., Ravichandran, N. K., Lindsay, L., and Broido, D., 2018, “Fermi Surface Nesting and Phonon Frequency Gap Drive Anomalous Thermal Transport,” Phys. Rev. Lett., **121**(17), p. 175901.

[181] Zeng, Z., Chen, C., Zhang, C., Zhang, Q., and Chen, Y., 2022, “Critical Phonon Frequency Renormalization and Dual Phonon Coexistence in Layered Ruddlesden-Popper Inorganic Perovskites,” Phys. Rev. B, **105**(18), pp. 1–7.

[182] Kirklin, S., Saal, J. E., Meredig, B., Thompson, A., Doak, J. W., Aykol, M., Rühl, S., and Wolverton, C., 2015, “The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies,” npj Comput. Mater., **1**(October).

[183] Wei, H., Bao, H., and Ruan, X., 2022, “Perspective: Predicting and Optimizing Thermal Transport Properties with Machine Learning Methods,” Energy AI, **8**(March), p. 100153.

[184] Bhattacharjee, D., Kundavu, K., Saraswat, D., Raghuvanshi, P. R., and Bhattacharya, A., 2022, “Thorough Descriptor Search to Machine Learn the Lattice Thermal Conductivity of Half-Heusler Compounds,” ACS Appl. Energy Mater., **5**(7), pp. 8913–8922.

[185] Fu, X., Wu, Z., Wang, W., Xie, T., Keten, S., Gomez-Bombarelli, R., and Jaakkola, T., 2022, “Forces Are Not Enough: Benchmark and Critical Evaluation for Machine Learning Force Fields with Molecular Simulations,” pp. 1–25.

[186] Tranås, R., Løvvik, O. M., Tomic, O., and Berland, K., 2022, “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., **202**(October 2021), p. 110938.

[187] Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B., 2022, “E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials,” Nat. Commun., **13**(1), pp. 1–11.

[188] Deng, B., Zhong, P., Jun, K., Han, K., Bartel, C. J., and Ceder, G., 2023, “CHGNet: Pretrained Universal Neural Network Potential for Charge-Informed Atomistic Modeling,” pp. 1–12.

[189] Jain, A., Veeravenkata, H. P., Godse, S., and Srivastava, Y., 2022, “High-Throughput Computational Discovery of 40 Ultralow Thermal Conductivity and 20 Highly Anisotropic Crystalline Materials.”

[190] Dronskowski, R., and Bloechl, P. E., 1993, “Crystal Orbital Hamilton Populations (COHP): Energy-Resolved Visualization of Chemical Bonding in Solids Based on Density-Functional Calculations,” J. Phys. Chem., **97**(33), pp. 8617–8624.

[191] Maintz, S., Deringer, V. L., Tchougréeff, A. L., and Dronskowski, R., 2016, “LOBSTER: A Tool to Extract Chemical Bonding from Plane-Wave Based DFT,” J. Comput. Chem., **37**(11), pp. 1030–1035.

[192] Yuan, K., Zhang, X., Chang, Z., Tang, D., and Hu, M., 2022, “Antibonding Induced Anharmonicity Leading to Ultralow Lattice Thermal Conductivity and Extraordinary Thermoelectric Performance in CsK2X (X = Sb, Bi),” J. Mater. Chem. C, pp. 15822–15832.

[193] Chang, Z., Ma, J., Yuan, K., Zheng, J., Wei, B., Al-Fahdi, M., Gao, Y., Zhang, X., Shao, H., Hu, M., and Tang, D., 2022, “Zintl Phase Compounds Mg3Sb2−xBix (x = 0, 1, and 2) Monolayers: Electronic, Phonon and Thermoelectric Properties From Ab Initio Calculations,” Front. Mech. Eng., **8**(April), pp. 1–11.

[194] Iotti, R. C., Ciancio, E., and Rossi, F., 2005, “Quantum Transport Theory for Semiconductor Nanostructures: A Density-Matrix Formulation,” Phys. Rev. B - Condens. Matter Mater. Phys., **72**(12), pp. 1–21.

[195] Rossi, F., and Kuhn, T., 2002, “Theory of Ultrafast Phenomena in Photoexcited Semiconductors,” Rev. Mod. Phys., **74**(3), pp. 895–950.

[196] Mukhopadhyay, S., Parker, D. S., Sales, B. C., Puretzky, A. A., McGuire,

M. A., and Lindsay, L., 2018, “Two-Channel Model for Ultralow Thermal Conductivity of Crystalline Tl3VSe4,” Science (80-.)., **360**(6396), pp. 1455–1458.

[197] Simoncelli, M., Marzari, N., and Mauri, F., 2022, “Wigner Formulation of Thermal Transport in Solids,” Phys. Rev. X, **12**(4), p. 41011.

[198] Zeng, Z., Zhang, C., Xia, Y., Fan, Z., Wolverton, C., and Chen, Y., 2021, “Nonperturbative Phonon Scatterings and the Two-Channel Thermal Transport in Tl3VSe4,” Phys. Rev. B, **103**(22), pp. 1–7.

Appendix A: Permission to Reprint

Permission for the text in Chapter 2 is authorized according to the following statement (link: https://journals.aps.org/copyrightFAQ.html#thesis):

As the author of an APS-published article, may I include my article or a portion of my article in my thesis or dissertation?

Yes, the author has the right to use the article or a portion of the article in a thesis or dissertation without requesting permission from APS, provided the bibliographic citation and the APS copyright credit line are given on the appropriate pages.

Permission for the text in Chapter 3 is authorized by ACS Publications by the “Policy on Theses and Dissertations” with the following statement (link: https://pubs.acs.org/pb-assets/acspubs/Migrated/dissertation-1632927826810.pdf):

Reprinted with permission from *ACS Appl. Mater. Interfaces* 2021, 13, 46, 55367–55379. Copyright 2021 American Chemical Society. You can view the full text of this article at http://pubs.acs.org/articlesonrequest/AOR-GCV3GEEXYR5CCRP6VDCG to download this article.

Permissions for the texts in Chapters 4 and 5 are authorized by the following statement:

Copyright on any research article in open access (OA) journals published by Springer Nature is retained by the author(s). Authors grant Springer Nature a license to publish the article and identify itself as the original publisher. Authors also grant any third party the right to use the article freely as long as its integrity is maintained and its original authors, citation details and publisher are identified.

OA articles in Springer Nature journals are predominantly published under Creative Commons Attribution v4.0 International licence (CC BY). For more information please read our journal policy.