



This is the accepted manuscript made available via CHORUS. The article has been published as:



Rapid prediction of phonon structure and properties using the atomistic line graph neural network (ALIGNN)

Ramya Gurunathan, Kamal Choudhary, and Francesca Tavazza

Phys. Rev. Materials **7**, 023803 — Published 24 February 2023

DOI: [10.1103/PhysRevMaterials.7.023803](https://doi.org/10.1103/PhysRevMaterials.7.023803)

Rapid Prediction of Phonon Structure and Properties using the Atomistic Line Graph Neural Network (ALIGNN)

Ramya Gurunathan ¹, Kamal Choudhary ^{1,2,3} and Francesca Tavazza ¹

¹*Materials Science and Engineering Division,
National Institute of Standards and Technology, Gaithersburg, 20899, MD, USA*

²*Theiss Research, La Jolla, 92037, CA, USA*

³*DeepMaterials LLC, Silver Spring, 20906, MD, USA*

Abstract

The phonon density-of-states (DOS) summarizes the lattice vibrational modes supported by a structure, and gives access to rich information about the material’s stability, thermodynamic constants, and thermal transport coefficients. Here, we present an atomistic line graph neural network (ALIGNN) model for the prediction of the phonon density of states and the derived thermal and thermodynamic properties. The model is trained on a database of over 14,000 phonon spectra included in the JARVIS-DFT (Joint Automated Repository for Various Integrated Simulations: Density Functional Theory) database. The model predictions are shown to capture the spectral features of the phonon density-of-states, effectively categorize dynamical stability, and lead to accurate predictions of DOS-derived thermal and thermodynamic properties, including heat capacity C_V , vibrational entropy S_{vib} , and the isotopic phonon scattering rate τ_1^{-1} . A comparison of room temperature thermodynamic property predictions reveals that the DOS-mediated ALIGNN model provides superior predictions when compared to a direct deep-learning prediction of these material properties as well as predictions based on analytic simplifications of the phonon DOS, including the Debye or Born-von Karman models. Finally, the ALIGNN model is used to predict the phonon spectra and properties for about 40,000 additional materials listed in the JARVIS-DFT database, which are validated as far as possible against other open-sourced high-throughput DFT phonon databases.

I. INTRODUCTION

The vibrational density of states (DOS) is a fundamental material feature, underpinning several properties related to thermodynamic stability and thermal conduction. Measuring the phonon density of states, for example by inelastic scattering tends to require access to synchrotron X-ray or high flux neutron sources, making high-throughput evaluations unfeasible[1]. Evaluation of computational phonon density-of-states via density functional theory (DFT) has become more mainstream via open-sourced software such as `phonopy`[2] or `almabte`[3], allowing for the formation of DFT-based phonon density-of-states databases[4, 5]. However, as this method requires evaluating the force sets between pairs of atoms, the calculation becomes increasingly expensive for complex unit cell materials.

For this reason, it is common to use simple, analytic approximations of the phonon

DOS when predicting thermal properties based on the Debye linear dispersion and Born-von Karman sinusoidal dispersion relations, for example[6–9]. However, these models can dramatically fail for materials that are highly anharmonic, have complex unit cells, or show large acoustic-optical band gaps[10, 11]. An attractive route to rapid predictions of phonon DOS and vibrational properties, directly from the crystal structure, is through deep learning[12]. In comparison to analytic approximations, the neural network phonon DOS offers a more accurate distribution of phonon modes including van Hove singularities and acoustic-optical band gaps, and it does not assume isotropy leading to false degeneracies between phonon modes. In our current work, anharmonic effects to the phonon DOS are not considered, as these require expensive training data in the form of higher-order force constants.

Crystal graph neural networks, which encodes features about atoms and their bonding environment in a non-Euclidean graph, have recently shown utility in predicting material properties such as formation energy, band gap, and elastic constants[13–16]. In their basic form, the nodes of a crystal graph represent individual atoms and the edges represent interatomic bonds. However, the nearest neighbor connectivity provides an incomplete picture of the local chemical environment, for example close packed bonding environments are difficult to distinguish, but incorporating information about bond angle distributions has provided a means to heuristically classify local structures[17]. The atomistic line graph neural network (ALIGNN) was developed recently as an extension of the crystal graph neural network, which explicitly incorporates bond angles by constructing a line graph over the original crystal graph representation. In the line graph, nodes correspond to bonds while edges correspond to pairs of bonds and therefore encode the bond angle cosine as a feature. By including bond connectivity and bond angle information, the ALIGNN model showed substantial improvements in prediction accuracy for formation energy, band gap, and magnetic moment when compared to a crystal graph convolutional neural network (CGCNN) [13], SchNet [18], MEGNet [19] and hand-crafted features such as classical force field inspired descriptors (CFID)[14]. The latter of which is not a graph representation, but centers on structural descriptors in the form of distribution functions such as the radial distribution or angular distribution functions. Numerous machine learning models have been developed to directly predict single, scalar-valued thermal and thermodynamic properties

such as heat capacity[20–22], vibrational entropy[21, 23], and thermal conductivity[24–27]. However, multiple-output prediction, as required for predicting the full phonon DOS spectrum, is relatively less developed[28, 29]. In computational spectroscopy, deep-learning forward models have been developed to predict a spectrum (e.g. infrared[30], X-ray[31, 32], photoemission[33]) directly from the crystal structure. However, as noted in Rankine *et al.*[32], these methods are less common than “reverse” mappings, which instead predict a property based on the measured spectrum.

In this work, we apply the ALIGNN model to predict the phonon density of states as well as derived thermodynamic and thermal properties, including the vibrational entropy, heat capacity, and phonon-isotope scattering rate. Recently, Kong *et al.* reported the Mat2Spec model for generating electronic and phononic DOS from material structure features[29]. In their work, both the input features (crystal structure) as well as the output features (DOS) are embedded as multivariate Gaussian distributions. The ALIGNN model, however, has been shown to perform well directly on discretized spectral training data when applied to the prediction of electric DOS with only modest improvements to model performance achieved using autoencoder-decoder segments to create a low-dimensional representation of the output features[34]. The work of Chen *et al.* demonstrated the feasibility of predicting phonon density of states using a graph neural network trained directly on discretized phonon DOS data. They applied a Euclidean neural network (E(3)NN), which uses a periodic crystal graph representation and convolutional filters made up of learned radial functions and spherical harmonics such that the neural network is equivariant to 3D group operations[35]. We build from these works with the explicit treatment of bond angles, training on the large dataset of DFT phonon density of states hosted on the JARVIS-DFT database[4, 5, 23], and by characterising the performance of the neural network on other structural and transport relevant properties of the DOS, including identification of negative phonon modes and prediction of DOS-derived material properties.

The paper first introduces the ALIGNN model and details of the JARVIS-DFT database used for training and validation. Next, the DOS-derived property equations are introduced to emphasize their differences in the weighting of phonon modes. Finally, we discuss the results of the model first in terms of the direct spectral features of the DOS and then in terms of the derived scalar material properties. We find that prediction of the DOS using

the ALIGNN model yields excellent results for the temperature-dependent heat capacity and vibrational entropy, and well as the phonon-isotope scattering rate.

II. METHODS AND THEORETICAL BACKGROUND

We will begin by introducing the atomistic line graph representation used to encode crystal structure in the neural network. Next, we will discuss the dataset used to train and validate the phonon DOS predictions. Finally, we will discuss the property models used to compute DOS-derived properties such as heat capacity, vibrational entropy, and phonon-isotope scattering rates.

A. Atomistic Line Graph Neural Networks

The open-sourced ALIGNN framework[14] is used here to encode crystal structure information in graph representations which then interface with a message passing neural network. This neural network framework can update the embeddings of nodes and edges in the graph while retaining graph connectivity and allowing neighboring nodes and edges to exchange information (pass messages) about their state. There are two graph representations used by the ALIGNN model, 1) an atomistic crystal graph in which nodes represent atoms and edges represent bonds, and 2) a line graph built from the crystal graph in which nodes represent bonds and edges represent bond pairs sharing a common atom. The crystal graph is represented as $G = (\nu, \epsilon)$, where ν is the set of nodes and ϵ is the set of edges, with a feature set inspired by the Crystal Graph Convolutional Neural Network (CGCNN) model[36]. The following 8 node features were used to describe individual atoms: electronegativity, covalent radius, group number, block, valence electron count, atomic volume, first ionization energy, and electron affinity[14]. The edge features, being associated with pairs of atoms/nodes (ν_i, ν_j) , are the interatomic bond distances r_{ij} . The bond distances are encoded as M -dimensional edge feature vectors using a radial basis function (RBF) expansion with support between 0 and 8 Å. We use 80 edge input features ($M = 80$) with 256 hidden features. Therefore, the edge feature set corresponding to the bond distance r_{ij} is represented as: $\epsilon_{ij} = \{\text{RBF}(\|r_{ij} - R_m\|)\}_{m=1}^M$. Here, R_m is the m^{th} center in the bond distance RBF support spanning $\mathbf{R} = [R_1, R_2, \dots, R_M]$.

Each edge in the crystal graph then becomes a node in the line graph. The line graph edges, denoted as t_{ijk} , corresponds to triplets of atoms, which in the atomistic graph are labelled by nodes ν_i, ν_j, ν_k and edges $\epsilon_{ij}, \epsilon_{jk}$. The line graph edge naturally represents a bond angle cosine spanned by the three atoms: $\theta_{ijk} = \arccos\left(\frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{jk}}{|\mathbf{r}_{ij}| |\mathbf{r}_{jk}|}\right)$, where the bolded notation \mathbf{r}_{ij} signifies the bond vector connecting nodes (ν_i, ν_j) . As before with the crystal graph edge features, the bond angle cosines are encoded using an RBF expansion such that the linegraph edge feature vector is defined as $t_{ijk} = \{\text{RBF}(\|\theta_{ijk} - T_m\|)\}_{m=1}^M$, where T_m is the m^{th} element in an array of bond angle RBF centers of length 40. The ALIGNN model can efficiently update atom and bond features by alternating message passing updates on both the crystal and line graphs. The crystalline materials treated in this work are represented using a periodic graph construction, expanded out to 20 nearest neighbors[14].

ALIGNN uses edge-gated graph convolution for updating nodes as well as edge features using a propagation function (f) for layer (l) and node (v_i) with associated feature vector (h_i) and neighbor list (N_i)[14]:

$$h_i^{(l+1)} = f(\nu_i^l \{\nu_j^l\}_{j \in N_i}) \quad (1)$$

As mentioned previously, for this work, we used 80 initial bond radial basis function (RBF) features, and 40 initial bond angle RBF features. The atom, bond, and bond angle feature embedding layers produce 64-dimensional inputs to the graph convolution layers. We used six ALIGNN update layers followed by six edge-gated graph convolution (each with hidden dimension of 256) updates on the bond line graph. Afterwards, the model performs a global average pooling of the final node vectors, which is used as input for fully connected regression layers that produce the final phonon DOS prediction. The ALIGNN model uses Sigmoidal Linear Unit (SiLU) activation layers because they are twice-differentiable[14], however this results in an output range of $(-0.5, \infty)$. In order to avoid any un-physical negative density-of-states values, we apply a final Rectified Linear Unit (ReLU) layer to the output tensors that replaces any small, negative values with 0, assuring only positive outputs.

Training was performed with a batch size of 64 samples, learning rate of 0.001, mean squared error loss function (MSE), AdamW optimizer, and 600 epochs. Further details about the ALIGNN model architecture and update procedure can be found in the original model reference[14], recent application to electronic structures[34], and in the shared code

repository on GitHub[37].

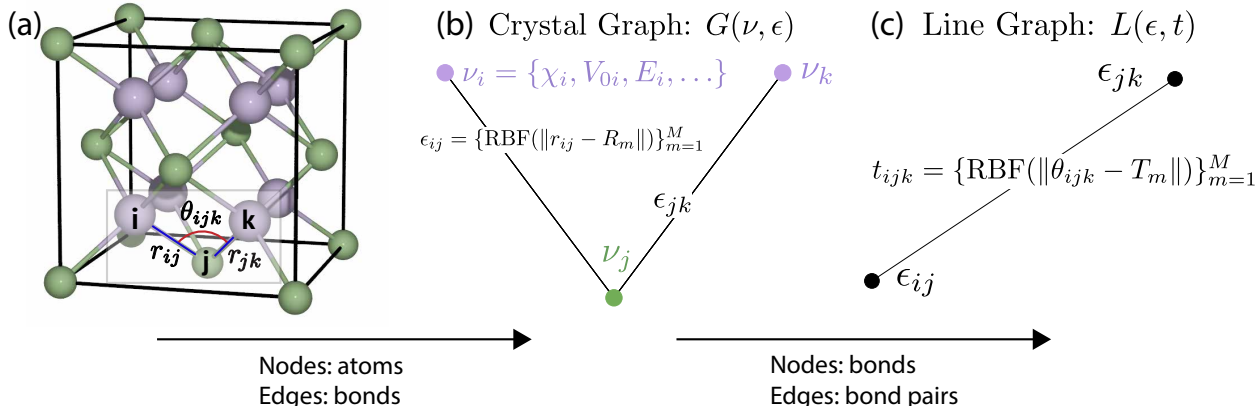


FIG. 1: Schematic of a crystal and line graph encoding of the Mg_2Si crystal structure. For simplicity of the undirected graph representations, these graphs are constructed by setting the maximum nearest neighbor value to 1. In the crystal graph (b), nodes represent atom sites and include an atomic feature set consisting of attributes like the electronegativity (χ), ionization energy (E), and volume per atom (V_{0i}). The edges in crystal graph represent bonds. Physically, the edge features represent bond distances r_{ij} which are encoded in the model using a radial basis function (RBF). The line graph (c) is constructed on top of the previous crystal graph, such that the nodes now represent the bonds of the crystal. The edges, therefore, represent pairs of bonds with a common atom or “triplets” featured by the bond angles, which once again are encoded using a radial basis function.

B. Training Dataset

The ALIGNN model is trained on a dataset of over 14,000 material-phonon DOS pairs computed at 0K using density functional theory (DFT), each labeled with a unique JARVIS-ID and hosted on the JARVIS-DFT public repository¹. The second-order force constants used to generate phonons were calculated using a finite difference method implemented in the VASP code[38] using the OptB88vdW functional[39], which includes non-local Van der Waals interactions. The k-point density and plane-wave energy cut-off were determined through an automated convergence method described in Choudhary and Tavazza[40], and the crystal structure optimization was performed with energy, forces, and stress relaxation[4]. A Brillouin zone integration was then performed to calculate the phonon DOS using Gaussian

¹ The DFT-computed electronic and phononic DOS hosted on JARVIS-DFT can be accessed using the `jarvis-tools` package or directly through the following `figshare` location: <https://doi.org/10.6084/m9.figshare.14745327>.

smearing interpolation in the `phonopy` package with a smearing width equal to 1/100 of the full phonon frequency range for the given material[41]. While the compounds in the phonon database are all evaluated to be energetically stable (i.e. negative formation energies evaluated at 0K), approximately 18% of the compounds in the JARVIS-DFT phonon dataset have over 10% of the integrated DOS in the negative phonon frequency range. These dynamically unstable compounds were included in the training set because predicting the existence of imaginary phonon modes is an important task in the vibrational analysis of a material. Therefore, the predictive range of the phonon deep-learning model should extend to negative frequencies.

The DOS was then discretized into bins of equal frequency width. We used a dataset with a frequency range of (-300 to 1000) cm^{-1} , as this is large enough to contain the minimum and maximum frequencies for nearly all of the compounds in the dataset. The bin size is 20 cm^{-1} , which is the bin size of the discretized DFT phonon DOS hosted on the JARVIS-DFT webpage. Additionally, the binned DOS is normalized by the maximum intensity such that the values range from 0 to 1, which improves numerical stability and quality of predictions. As stated in the work of Chen *et al.* on the e3nn phonon DOS model, the correctly-scaled DOS is easily recovered from the normalized version because of the physical definition requiring that the integrated DOS equal $3N$, where N is the number of atoms in the formula unit[8, 29]. This equality provides a straightforward route to recover the appropriately scaled DOS spectrum such that comparisons between different phonon DOS are meaningful.

During training, the dataset was randomly partitioned into a 80 %-10 %-10 % training-validation(during training)-test(fully blind) split. The attributes of the training set are highlighted in Figure 2. Oxygen is by far the most abundant element in this dataset, which consists mainly of binary and ternary compounds in the cubic and tetragonal structure types. However, all seven crystal systems are represented in the training dataset.

C. Property Models

We focus on the evaluation of three thermodynamic and thermal properties based on the phonon density of states ($g(\omega)$). The first is the harmonic contribution to the heat capacity,

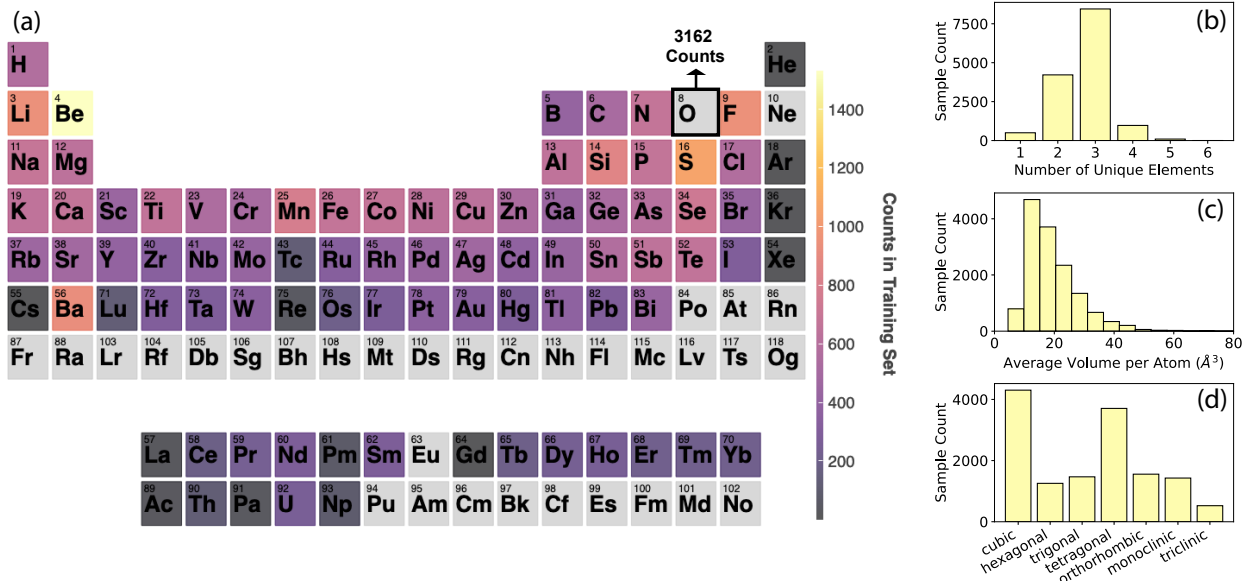


FIG. 2: Attributes of the phonon density-of-states training data set comprising 11384 total materials. Panel (a) shows the enrichment of each element in the training set. Oxygen was removed from the colormap because it is highly enriched with 3162 total counts. The histograms in panels (b)-(d) indicate that the training set is enriched with binary and ternary compounds in the cubic or tetragonal crystal system with an average volume per atom of about (10 to 20) \AA^3 .

a measure of the heat stored by the phonon modes of a material, which tends to be the majority contribution to the overall heat capacity. Within this harmonic approximation, the lattice does not undergo thermal expansion, and so it is natural to define the heat capacity at constant volume[42]. The heat capacity can be determined directly from the phonon density of states, where the phonon modes are weighted by their energy and the temperature derivative of the Bose-Einstein distribution[43].

$$C_V = \int k_B (\hbar\omega/k_B T)^2 \frac{\exp(\hbar\omega/k_B T)}{(\exp(\hbar\omega/k_B T) - 1)^2} g(\omega) d\omega \quad (2)$$

At high temperature, when the full vibrational spectrum is excited, the heat capacity will approach a thermodynamic limit of $1 k_B$ per phonon mode, yielding the Dulong-Petit limit for molar heat capacity of $C_V = 3NR$, where N is the number of atoms per formula unit and R is the gas constant. The heat capacity begins to saturate at the Dulong-Petit limit near the Debye temperature (T_D), or the temperature at which the highest phonon mode is activated (see Figure 3). The Debye temperature relates strongly to the stiffness of the

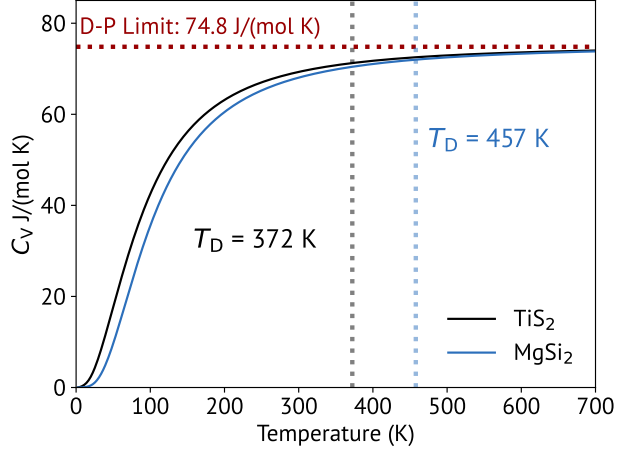


FIG. 3: Comparison of molar heat capacity versus temperature for two materials (TiS_2 and MgSi_2) with the same Dulong-Petit limit, but differing Debye temperatures (T_D). As shown, the heat capacity values approach the Dulong-Petit limit and therefore start to converge above the Debye temperature.

material[42]. As a result, we examine two types of heat capacity datasets: 1) C_V calculated at a constant temperature (i.e., 300 K) and 2) C_V calculated at a fixed fraction of the Debye temperature (i.e. $0.5T_D$). The constant temperature dataset may be useful when comparing materials for a given application. However, the constant fraction of the Debye temperature is useful when assessing the accuracy of the model for a similar weighting of the density of states in the Equation 2 integrand. Note that by treating the temperature dependence using an analytic expression based on the phonon DOS, we avoid having to train separate neural networks for each desired temperature.

The vibrational entropy S_{vib} is also evaluated here and describes the range of momentum and position coordinates probed by atoms as they vibrate in a material. As temperature increases, more phonons are excited and atoms vibrate at higher amplitudes, so the vibrational entropy contribution should increase in magnitude. The vibrational entropy is known to have a role in polymorphic phase transitions and can stabilize lower-symmetry structures with longer bond lengths[44, 45]. Additionally, vibrational entropy can significantly influence solubility limits and the location of phase boundaries, making it important to quantify[46]. The vibrational entropy can be determined from a different weighting of the density of states, which stems from multiplying the partition functions for the $3N$ oscillators available in the material. Lower frequency phonons tend to be weighted more heavily in

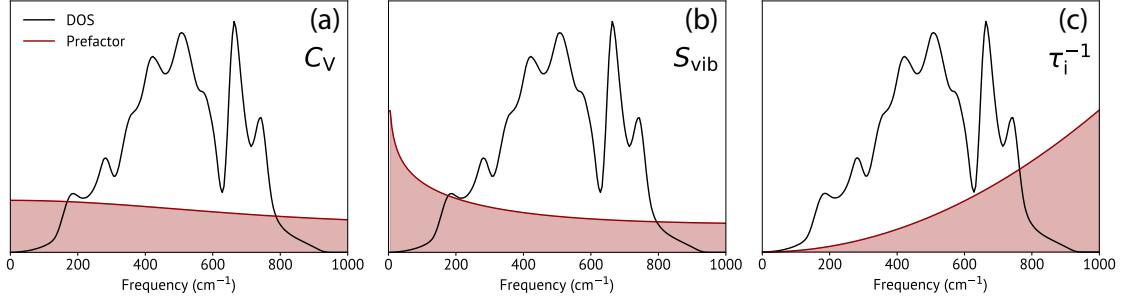


FIG. 4: DOS-derived thermodynamic and thermal properties will weigh regions of the phonon spectrum differently, as portrayed in these schematics. (a) The heat capacity weighting is relatively constant with phonon frequency, while (b) the vibrational entropy weighs low-frequency phonon modes more heavily. Finally, (c) the isotope-phonon scattering rate is proportional to $\omega^2 g(\omega)$ and therefore weighs high-frequency phonons more heavily.

S_{vib} , and therefore typically require high-accuracy descriptions (see Figure 4²).

$$S_{\text{vib}} = \int k_{\text{B}}[(n(\omega) + 1)\ln(n(\omega) + 1) - n(\omega)\ln n(\omega)]g(\omega)d\omega; \text{ where } n(\omega) = (\exp(\hbar\omega/k_{\text{B}}T))^{-1} \quad (3)$$

Finally, the $g(\omega)$ plays an important role in describing elastic phonon scattering processes as it defines the scattering phase space or the set of states that an incident phonon can scatter into. To exemplify its role in scattering problems, we calculate the phonon scattering rate due to phonon-isotope interactions (τ_{i}^{-1}), using the natural isotopic abundance for the given material[47, 48].

$$\tau_{\text{i}}^{-1} = \int \frac{\pi}{6} V_{\text{at}} \Gamma \omega^2 g(\omega) d\omega \quad (4)$$

Here, V_{at} is the volume per atom and Γ is the mass variance introduced to the lattice by isotopes (see [49] and [48] for details). As indicated by the factor of ω^2 in Equation 4, higher frequency phonons will be more heavily weighted in the scattering rate calculation (see Figure 4).

² In Figure 4, the example spectrum is that of Al_2O_3 (JARVIS-ID: JVASP-32) and can be obtained using the `jarvis-tools` Python package. Note that this example spectrum comes from a phonon DOS dataset with the following specifications: frequency range of (0 to 1000) cm^{-1} and bin size of 5 cm^{-1} . These binning parameters differ from that of the spectra in our training dataset.

III. RESULTS AND DISCUSSIONS

The model predictions are discussed in this section, first in the context of the direct phonon spectrum prediction and then in terms of derived properties, including the temperature-dependent heat capacity C_V , vibrational entropy S_{vib} , and the phonon-isotope scattering rate τ_i .

A. Model Performance

We use the average mean absolute error (MAE) across the binned density-of-states to evaluate the performance in predicting the direct spectrum. The average MAE is defined as

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (5)$$

where \hat{y}_i is the predicted value of the i^{th} bin, y_i is the target value of the i^{th} bin, and n is the total number of bins.

Figure 5 summarizes the MAE distribution and trends in the test set. The samples of the test set are concentrated at lower MAE values with 78 % of the samples showing an MAE of less than 0.086, as represented in the histogram in Figure 5a. To better interpret the MAE values, we show examples of the DFT versus ALIGNN-predicted spectra from the first eight MAE bins, which comprise 99.4 % of the test set. To better convey the model performance, we show additional example spectra in Supplementary Section S1[50] at finer intervals of MAE. Starting around an MAE level of 0.05, the ALIGNN model begins to miss or average over certain peaks in the spectrum, which is an error seen in deep-learning models for computational spectroscopy[30, 33]. The model does, however, tend to capture peak placement and the overall frequency range of the material’s phonon DOS relatively well.

Interestingly, we find that MAE is inversely correlated with the average atomic volume in the compound, likely due to the fact that a smaller atomic volume tends yield a higher maximum frequency in the phonon DOS. As a result, the model needs to predict peaks across a larger frequency range in these compounds. We additionally note that the MAE tends to increase with the number of unique elements in the compound as well as reduced

symmetry of the crystal system. Increased complexity of the formula unit will yield more numerous vibrational modes while reduced symmetry results in degeneracy-breaking, all of which tend to yield a greater number of peaks in the density-of-states. Therefore, these trends are logical, but may additionally reflect biases in the dataset, which is enriched with binary and ternary compounds as well as cubic and tetragonal materials.

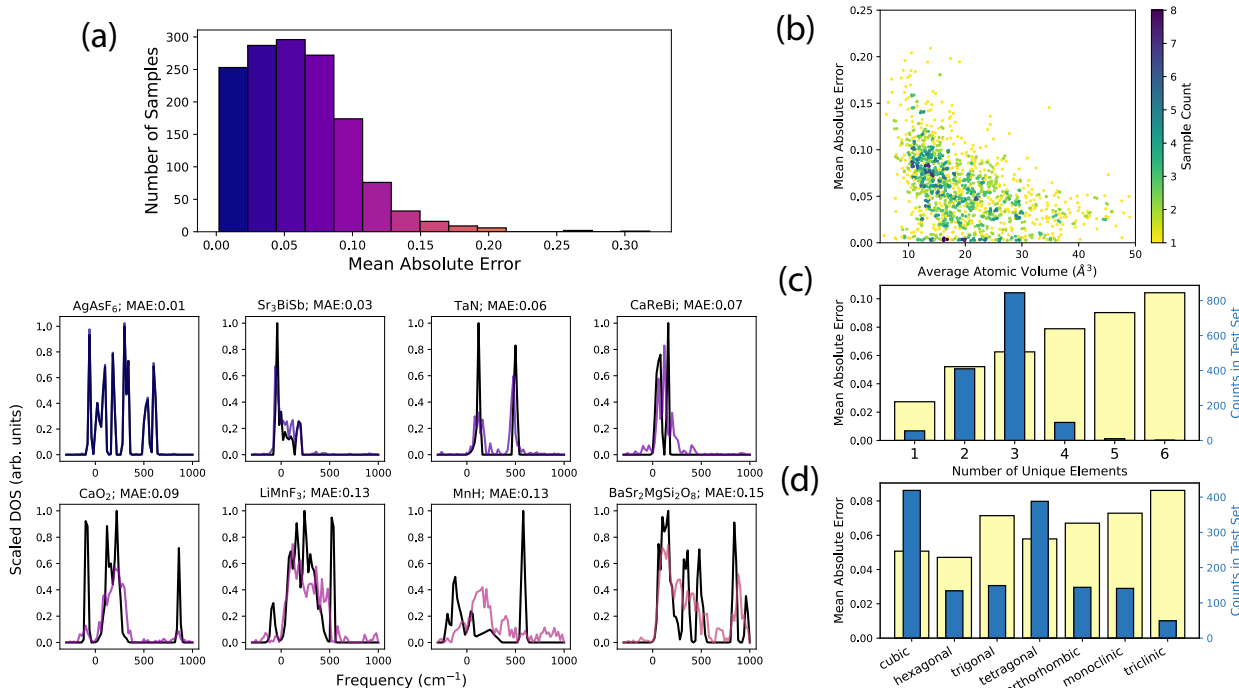


FIG. 5: Assessment of model performance in terms of the mean absolute error (MAE) of individual sample DOS spectra. Panel (a) shows the MAE histogram for ALIGNN-predicted phonon DOS in the test set, which highlights the concentration of samples at low MAE. Below, we show example spectra comparing the DFT DOS (black) to the ALIGNN DOS (colored) from the first 8 bins, which comprise 99.4% of the test set. Panel (b) indicates that the MAE tends to decrease with increasing volume per atom. While panel (c) may suggest that the MAE grows with increasing number of elements in the compound, although this trend is confounded by the fact that the test set is highly enriched with binary and ternary compounds. Finally, panel (d) suggests a more complicated relationship with crystal system, where MAE is anti-correlated with both symmetry and abundance in the training set.

In order to compute meaningful DOS-derived properties, we must first filter out dynamically unstable compounds from the dataset, or compounds with imaginary or negative frequency phonon modes. Note that because our training data consists entirely of phonon structure computed at 0K, this classification of dynamical stability is strictly for 0K and does not

apply at any arbitrary temperature. We chose a more tolerant definition of dynamical stability in order to retain as many samples as possible, and because imaginary phonon modes are often used to interpret structural or metal-insulator phase transitions[51, 52]. If the integrated area below 0 cm^{-1} composed less than 10 % of the total integrated DOS, then the sample was labelled dynamically stable. The confusion matrix (Figure 6) is used to compare the ALIGNN and DFT phonon DOS predictions in terms of this dynamical stability criteria. Note, however, that the ALIGNN model was not re-trained as a classifier model for this task. The ALIGNN model is being used as a regression model to fit the full phonon DOS, and then the 10 % integrated DOS rule is used to classify the spectra. While it is important to note that the dataset is overwhelmingly composed of dynamically stable compounds, the ALIGNN model classifies dynamical stability with 90 % accuracy, where the mostly likely prediction error is a false “stable” labelling. The precision of the classification (i.e., percentage of predicted stable compounds which are DFT stable) is 92.4 %, while the recall (i.e., percentage of DFT stable compounds correctly predicted as stable) is 95.2 %.

B. Derived Thermal Property Predictions

We then analyze model performance in terms of DOS-derived properties, including the molar heat capacity C_V (in J/(mol K)), the molar vibrational entropy S_{vib} (in J/(mol K)), and the phonon-isotope scattering rate (in GHz). The relationship between properties generated from the target (DFT) DOS versus the predicted (ALIGNN model) DOS are depicted as scatter plots in Figure 7 with scatter points representing samples in the test set. The central dashed line represents the 1:1 correlation while the surrounding dotted lines bound the width of the interquartile range for the target property distribution, as a way to quantify the spread of the samples.

The molar heat capacity at 300 K (Figure 7b) shows a concentration of samples at intervals of $3NR$ since several samples have reached the Dulong-Petit limit for phononic heat capacity by room temperature. As this greatly simplifies the distribution of heat capacity, the correlation coefficient R^2 between the target and predicted C_V values is 0.998 with a mean absolute error to mean absolute deviation ratio (MAE:MAD) of 0.03, indicating a very low error prediction with respect to the spread in the property distribution (see Figures

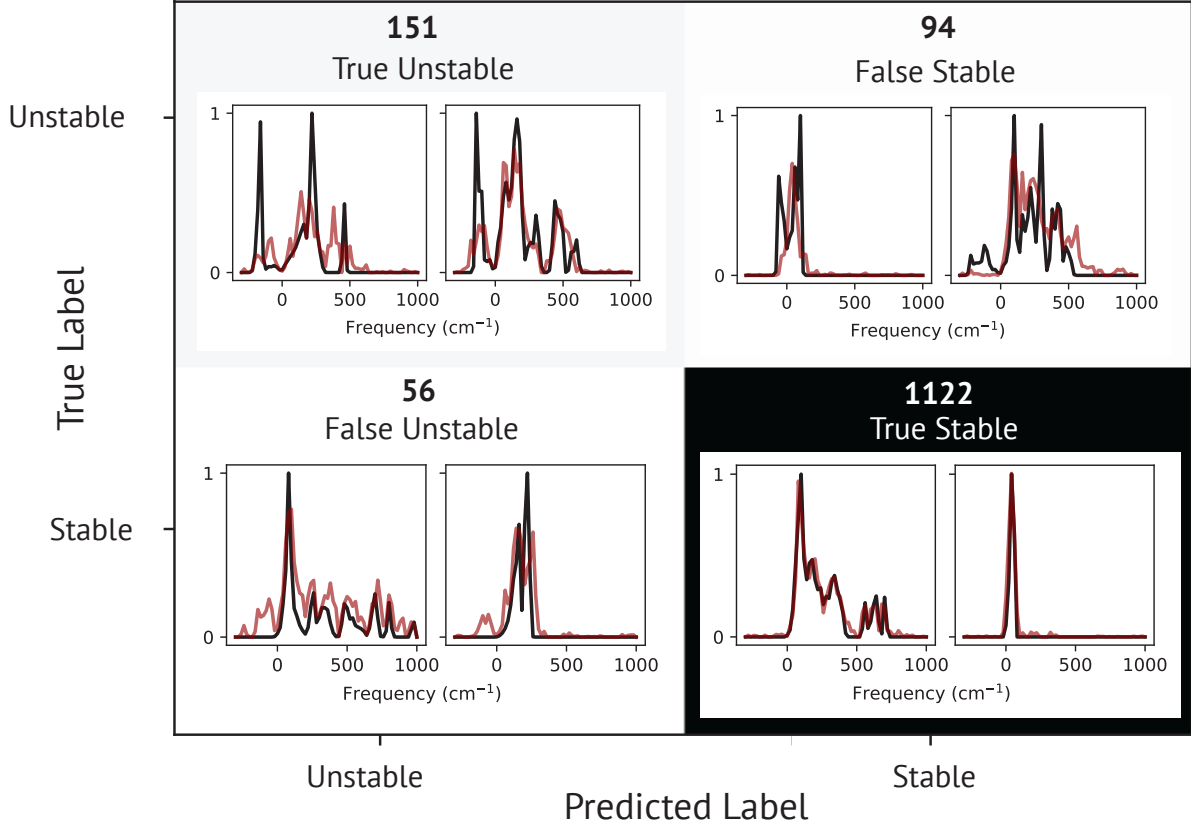


FIG. 6: Both the DFT and ALIGNN-predicted phonon DOS are classified as dynamically unstable if more than 10 % of the integrated DOS is in the negative frequency range. Based on this classification rule, the confusion matrix compares the DFT prediction (true label) to the ALIGNN prediction (predicted label) for all samples in the test set. As depicted, the majority of sample are correctly classified, but the test set contains primarily dynamically stable compounds. Randomly selected example spectra are shown for each category in the confusion matrix to illustrate the types of DOS errors that can lead to misclassification. The DFT spectrum is shown in black while the ALIGNN-predicted spectrum is shown in red.

8 and 7 for summary of DOS-derived property error metrics). The mean absolute deviation (MAD) is the average distance between each data point (y_i) and the mean (\bar{y}) of the dataset:
$$\text{MAD} = (1/n) \sum_{i=1}^n |y_i - \bar{y}_i|.$$

To avoid the influence of the Dulong-Petit limit in the model evaluation, we additionally compare C_V values computed at half of the Debye temperature (T_D) for the entire test set. As shown in Figure 7b, this dataset is more distributed since no samples have reached their Dulong-Petit limit of $3NR$. Even without the simplification imposed by the Dulong-Petit

limit, there is only a modest reduction in model performance in the $0.5T_D$ dataset. Figure 7i shows the MAE in the C_V prediction varying within 2 J/(mol K) over a wide percentage range of the Debye temperature.

The molar vibrational entropy S_{vib} also shows a robust correlation between target and predicted values ($R^2 = 0.986$), but larger error to spread ratio (MAE:MAD = 0.1). As discussed in Section II, one reason for the greater prediction error may be that the vibrational entropy integration heavily weights low-frequency phonons. In Supplementary Section S3, we show the DFT versus ALIGNN phonon DOS for three samples with the highest residuals in S_{vib} . A common feature of these spectra is a large DOS peak near 0 cm^{-1} , which when improperly predicted yields significant S_{vib} prediction error. Moreover, unlike the C_V prediction, for which the MAE peaks around 25 % of the Debye temperature, the MAE for S_{vib} steadily increases with temperature and saturates above the Debye temperature. Lastly, in the phonon-isotope scattering rate calculation (τ_i^{-1}), compounds with no natural isotopic abundance and therefore no phonon-isotope scattering were filtered out. Nonetheless, the property distribution was highly right skewed. Once again, the ALIGNN DOS model performed well yielding an MAE:MAD ratio of 0.11. Notably, in this case, some of the variance between samples stems not from the phonon DOS, but the mass variance factor Γ , which is calculated separately in the analytic expression. The prediction of the properties of interest (C_V , S_{vib} , and τ_i^{-1}) is robust against several of the errors in the direct DOS prediction apparent from visual inspection. The quality of the property prediction remains high because they are computed as the integral of the DOS, weighted by a frequency-dependent prefactor. Therefore, for these derived properties, it is important for the ALIGNN model to place the phonon peaks in the correct frequency range, but these results are robust against many distortions in the shape of the peak. As shown in Figure 4, the S_{vib} prefactor preferentially weights low-frequency phonons while the τ_i^{-1} prefactor weights high-frequency phonons, as such, DOS errors in these regimes will more greatly impact the corresponding property. In Supplementary Figure S2, we illustrate the types of DOS prediction errors that lead to large errors for each DOS-derived property.

To summarize our approach, here we apply the ALIGNN deep learning model to generate the phonon DOS, and subsequently apply the equations in Section II to derive various thermal and thermodynamic properties. We can compare this DOS-mediated approach to

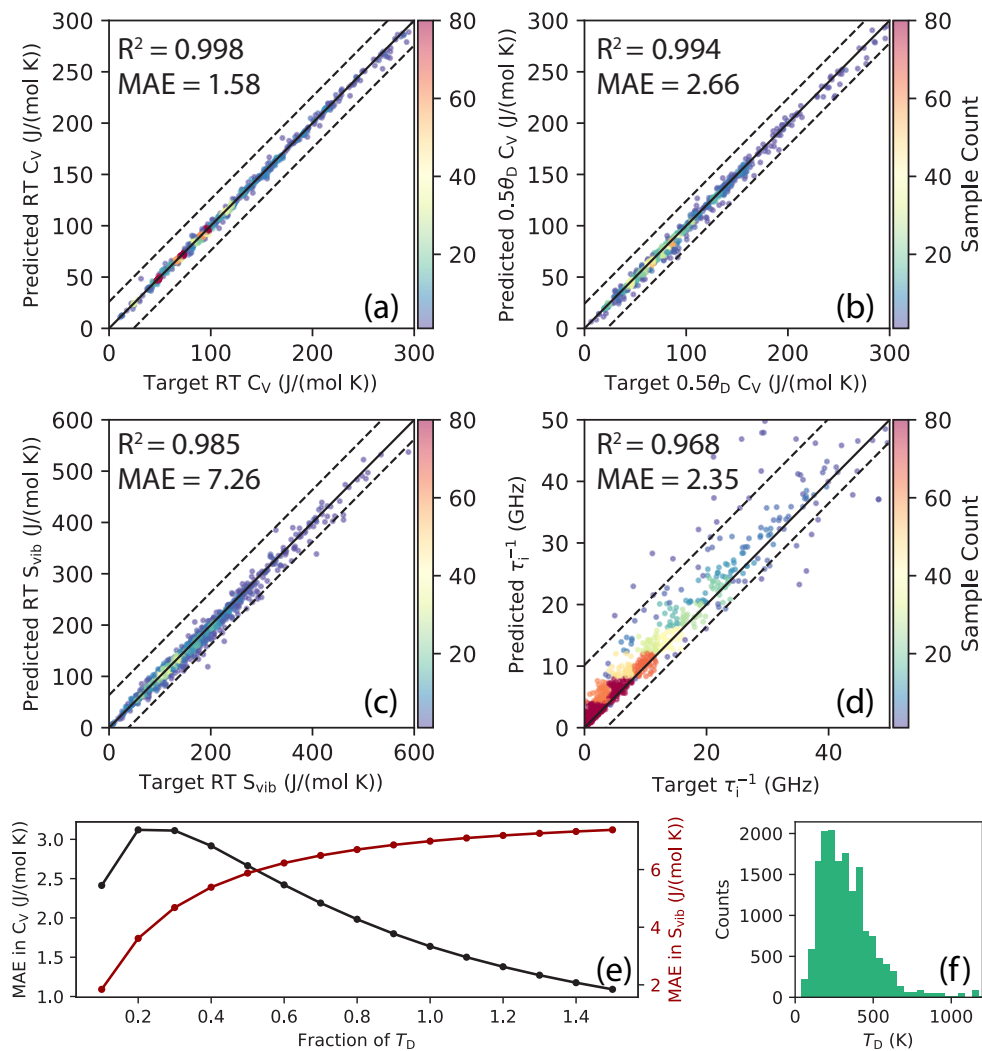


FIG. 7: Properties derived from the target (DFT) versus the predicted (ALIGNN) phonon DOS. Note that only samples determined to be dynamically stable both by DFT and ALIGNN are included in thermal property assessments. In panels (a-d) scatter points represent individual samples in the test set, while the center dashed line shows the 1:1 correlation and the dotted lines highlight the width of the interquartile range in the distribution of target values. The heat map values come from a 2D histogram showing the distribution of samples, where red regions indicate a large concentration of samples. Panels (a) and (b) both depict the molar heat capacity C_V , at 300 K and 50 % of each material’s Debye temperature, respectively. At 300 K, heat capacity values are closely clustered around intervals of $3NR$ since several samples have reached the Dulong-Petit limit. Since this physical limit greatly simplifies the prediction, we also show that the C_V predictions remain accurate at across intervals of the Debye temperature, as shown in the plot of the mean absolute error (MAE) for the C_V prediction as a function of fractional Debye temperature (e). Panel (c) depicts the room temperature (RT) molar vibrational entropy S_{vib} , and shows a strong trend between predicted and target values with a few instances of under-predictions. Panel (d) depicts the phonon-isotope scattering rate. The natural isotope abundance for each material was used, which was attained from the isotope database in the `phonopy` package[2]. Compounds without any known isotope variation were excluded. Finally, the distribution of Debye temperatures for the samples in the test set is shown in panel (f).

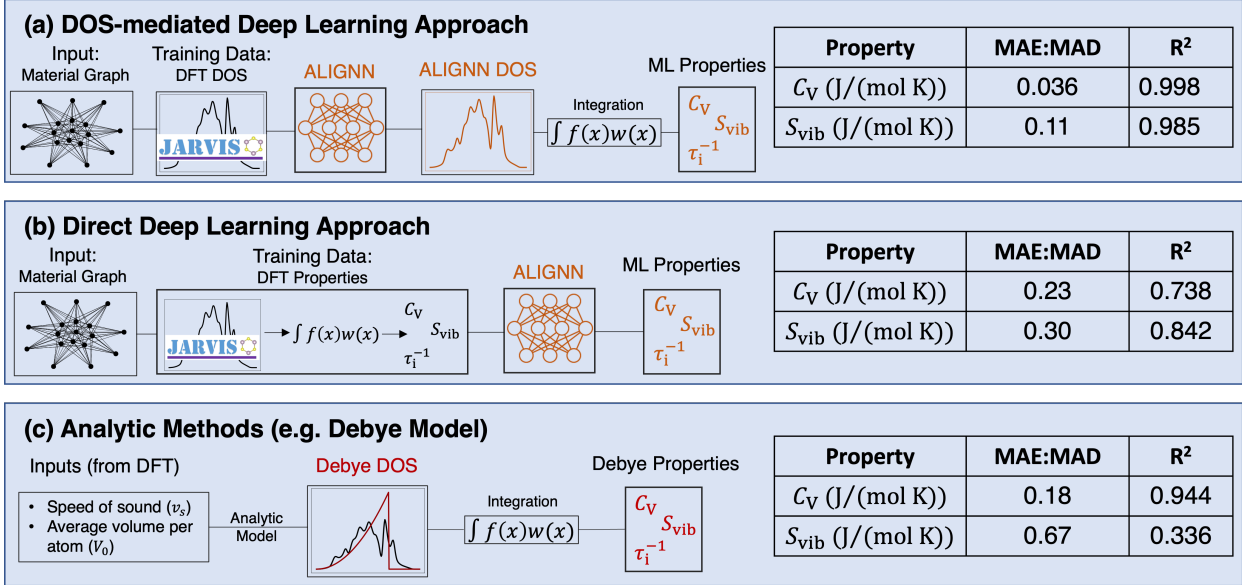


FIG. 8: Model performance is compared for three techniques used to compute DOS-derived material properties, showing that the DOS-mediated approach emphasized in this work yields the most accurate prediction. Techniques (a) and (b) differ primarily in the placement of the deep learning step relative to the DOS integration step used to derive the material properties. In panel (c) we show the results for the best analytic DOS approximation trialled, the Debye approximation. Error metrics include the ratio of the mean absolute error to the mean absolute deviation (MAD) of the target property distribution (MAE:MAD) as well as the correlation coefficient between target and predicted values (R^2). Both the molar heat capacity and vibrational entropy are calculated here at room temperature (RT; 300 K).

a direct prediction of either heat capacity or vibrational entropy using an ALIGNN model trained directly on the property values. To perform this calculation, we compile two training datasets, one comprising DFT heat capacity values and the other comprising DFT vibrational entropy values (i.e., C_V or S_{vib} computed from the DFT phonon DOS using Equations 2 and 3). We then train separate ALIGNN property models using each dataset. In order to replicate the ALIGNN phonon DOS model as closely as possible in these follow-up tests, we use the same training-validation-test split of the JARVIS-IDs in the database, and we use the same configuration settings related to the graph representation of the material (e.g. maximum number of nearest neighbors, node feature set) and the training (e.g. learning rate, optimizer).

A major advantage of using the DOS-mediated approach is that the temperature dependence of the C_V and S_{vib} is embedded in the model. By modelling the full vibrational structure,

we provide knowledge of which phonon modes can be excited at a given temperature. In contrast, a direct deep learning prediction of temperature-dependent C_V and S_{vib} may require either re-training the neural network for each desired temperature or accumulating property training data labelled with temperature that spans the temperature range of interest. Kauwe *et al.*[20] comments on the difficulty of developing empirical models[53, 54] or direct machine learning models for temperature-dependent heat capacity because thermochemical property data as a function of temperature is not easily attained. Additionally, composition-based features cannot adequately capture nuances regarding the active vibrational modes at a given temperature. When comparing room temperature properties, we find that the DOS-mediated approach yields a significantly lower model prediction error versus the direct deep-learning approach (see Figure 8). The MAE for the direct ALIGNN prediction is 9.60 J/(mol K) and 16.9 J/(mol K) for the RT C_V and S_{vib} , respectively, compared to the DOS-mediated approach values of 1.58 J/(mol K) and 7.26 J/(mol K). The especially poor direct ALIGNN prediction of the heat capacity comes from the concentration of samples around discrete C_V values, as the ALIGNN model tends to produce a smoother property distribution. These results reaffirm that predicting the phonon structure from crystal structure using deep learning is preferable to a direct machine learning prediction of phonon-based properties.

To better understand the performance of the ALIGNN model, we compared these DOS-derived properties to predictions from analytic approximations of the phonon density-of-states using both the Debye (linear) and Born-von Karman (sinusoidal) approximations of the ω vs. q relation [9]. Although these dispersion approximations have known limitations, they are still frequently applied to rapidly predict heat capacity or phonon scattering rates[7, 8, 55, 56]. Example approximations for the phonon dispersions and phonon DOS are depicted in Figure 9. We confirm that for room temperature property predictions, the Debye and Born-von Karman approximations yield much lower model performance. The MAE for the RT molar C_V is 7.90 (Debye) and 10.5 J/(mol K) (Born-von Karman), while the MAE for the RT molar S_{vib} is 30.0 (Debye) and 43.9 J/(mol K) (Born-von Karman).

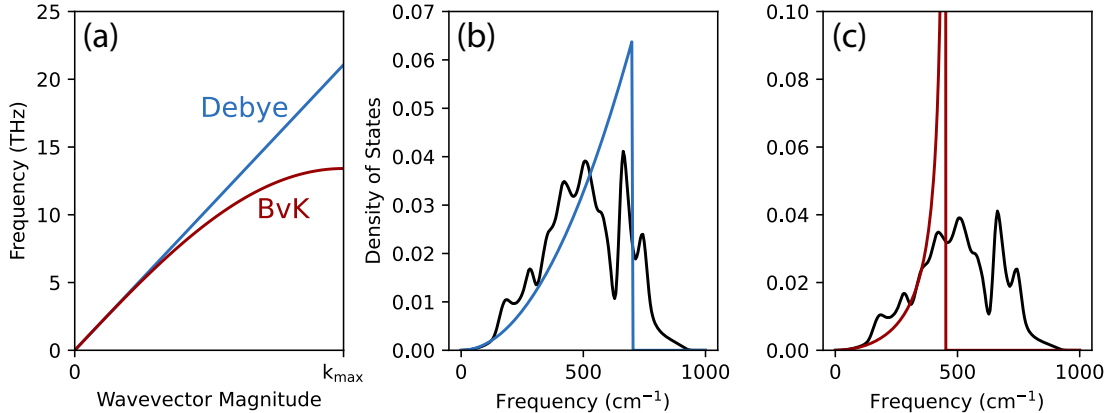


FIG. 9: Schematic of the Debye (linear) and Born-von Karman (sinusoidal) approximations of the phonon dispersion is shown in panel (a). Panels (b)-(c) depict the resulting density of states approximation for the example case of Al_2O_3 compared to the DFT density of states.

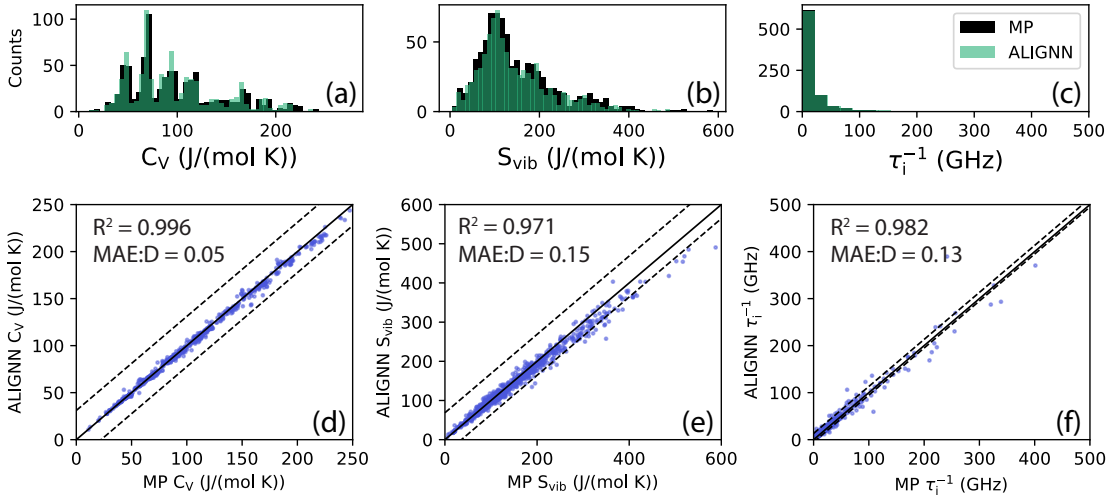


FIG. 10: Comparison between the ALIGNN thermal property predictions and those calculated from the Materials Project density functional perturbation theory (DFPT) phonon structure database. We found 830 overlapping compounds and show a close correspondence between the DFPT and deep-learning results. The R^2 coefficients and MAE to MAD ratio (abbreviated as MAE:D) are shown on the subplots. Note that this relationship holds even though the phonon DOS used for training of the ALIGNN model were generated using a finite-difference rather than a perturbative approach.

C. Large-Scale Prediction on New Compounds

We then applied the trained ALIGNN model to 41,000 crystal structures in the JARVIS-DFT database with unknown phonon DOS and vibrational properties. In Supplementary Section S2, we list the 10 highest and lowest predicted property values and corresponding

compositions for each of the three material properties considered here: specific heat capacity at 300 K C_V (in J/(kg K)), specific vibrational entropy at 300 K S_{vib} (in J/(kg K)), and phonon-isotope scattering rate τ_i^{-1} (in GHz). The properties showed expected trends: the specific heat capacity and vibrational entropy mainly showed an inverse correlation with molar mass and bond strength. The compounds with the highest isotope scattering rate were dominated by light elements with large isotopic variation (such as B and Ge). In Figure 11 we show the ALIGNN-predicted phonon DOS yielding the lowest and highest contribution to the heat capacity, vibrational entropy, and phonon-isotope scattering rate. The heat capacity and vibrational entropy comparisons were made on a “per oscillator” basis through normalizing by $3N$ to emphasize the influence of phonon DOS shape on these properties. As a general trend, a shift of phonon modes from high to low frequency yields higher room temperature heat capacity and vibrational entropy, but lower phonon-isotope scattering rate.

Additionally, about 830 of the predicted compounds overlapped with the entries in the Materials Project (MP) phonon database generated using density functional perturbation theory [5]. We observe a very close correspondence between our ALIGNN-predicted thermal properties and those derived from the MP phonon spectra (see Figure 10) even though the calculation method for the phonon dataset used in training of the neural network relied on the finite-difference method rather than perturbation theory.

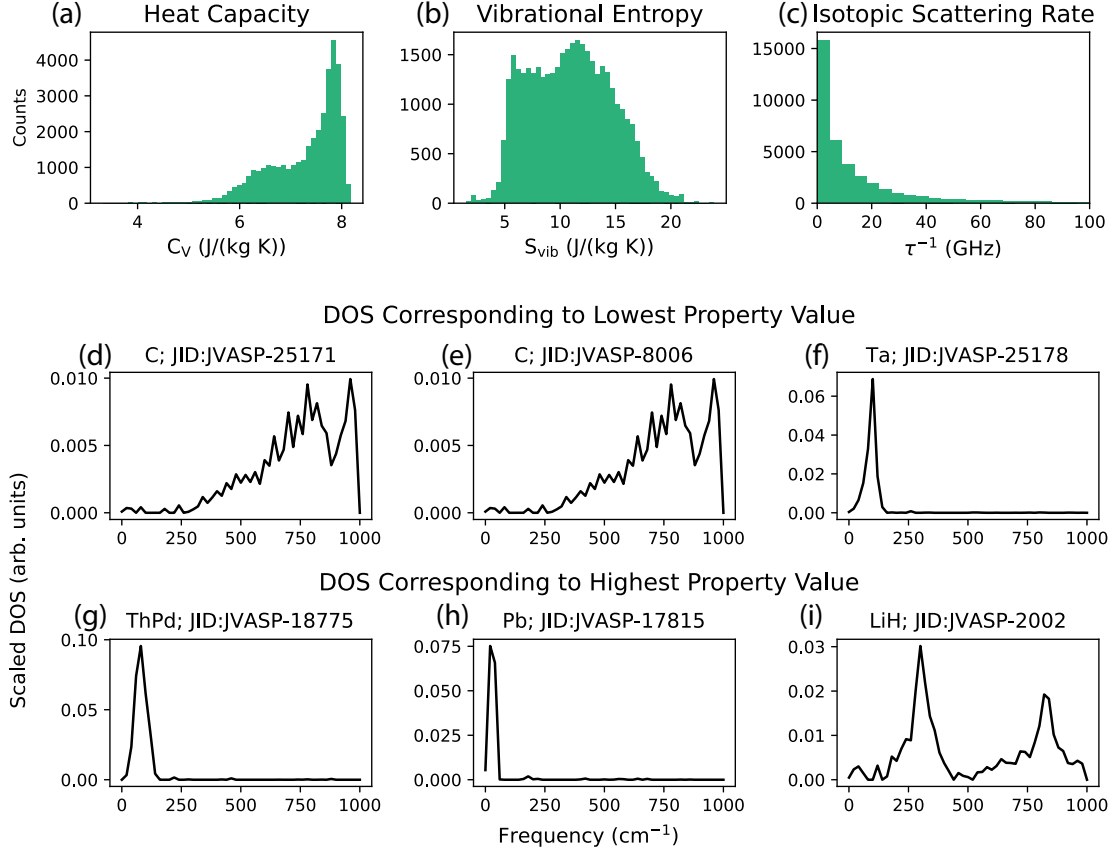


FIG. 11: Summary of predicted phonon properties for 41,000 previously unlabeled compounds in the JARVIS-DFT database. Panels (a-c) show the distribution of the 300 K gravimetric C_V and S_{vib} as well as the τ_i^{-1} . Panels (d-f) show the ALIGNN-predicted spectra yielding the lowest C_V , S_{vib} , and τ_i^{-1} contribution. Note that in the case of C_V and S_{vib} , this is the lowest contribution on a “per-oscillator” basis (i.e. comparing C_V and S_{vib} values divided by $3N$). Similarly, panels (g-i) show the ALIGNN-predicted spectra yielding the highest contribution to the DOS-derived properties.

IV. CONCLUSION

The atomistic line graph materials representation preserves the connectivity of the crystal structure and explicitly encodes features describing the atoms, bonds, and bond angles. The technique is shown here to give reasonable predictions of the vibrational structure and properties of a material. In particular, the ALIGNN DOS model yields excellent predictions for three thermodynamic and thermal properties studied, the heat capacity, the vibrational entropy, and the phonon-isotope scattering rate. Through a comparison of room-temperature thermodynamic property predictions, we see that the placement of the deep-learning step relative to the physics-based integration step makes a large difference in prediction quality.

We find that using the deep-learning ALIGNN model to predict the DOS spectrum is preferable to learning the DOS-derived properties directly, yielding both better accuracy and rich information in one deep-learning step, since the DOS encodes for numerous material properties and their temperature dependence.

V. ACKNOWLEDGEMENTS

The authors thank the National Institute of Standards and Technology for funding, computational, and data management resources. R.G. acknowledges funding through the NRC Research Associateship Program. Contributions from K.C. were supported by the financial assistance award 70NANB19H117 from the U.S. Department of Commerce, National Institute of Standards and Technology.

-
- [1] R. Hanus, R. Gurunathan, L. Lindsay, M. T. Agne, J. Shi, S. Graham, and G. Jeffrey Snyder, *Applied Physics Reviews* **8**, 031311 (2021), <https://doi.org/10.1063/5.0055593>.
 - [2] A. Togo and I. Tanaka, *Scripta Materialia* **108**, 1 (2015).
 - [3] J. Carrete, B. Vermeersch, A. Katre, A. van Roekeghem, T. Wang, G. K. Madsen, and N. Mingo, *Computer Physics Communications* **220**, 351 (2019).
 - [4] K. Choudhary, K. F. Garrity, A. C. E. Reid, B. DeCost, A. J. Biacchi, A. R. Hight Walker, Z. Trautt, J. Hatrick-Simpers, A. G. Kusne, A. Centrone, A. Davydov, J. Jiang, R. Pachter, G. Cheon, E. Reed, A. Agrawal, X. Qian, V. Sharma, H. Zhuang, S. V. Kalinin, B. G. Sumpter, G. Pilania, P. Acar, S. Mandal, K. Haule, D. Vanderbilt, K. Rabe, and F. Tavazza, *npj Computational Materials* **6**, 173 (2020).
 - [5] G. Petretto, S. Dwaraknath, H. P. Miranda, D. Winston, M. Giantomassi, M. J. Van Setten, X. Gonze, K. A. Persson, G. Hautier, and G. M. Rignanese, *Scientific Data* **5**, 1 (2018).
 - [6] T. A. R. Purcell, M. Scheffler, L. M. Ghiringhelli, and C. Carbogno, *Accelerating materials-space exploration by mapping materials properties via artificial intelligence: The case of the lattice thermal conductivity* (2022).
 - [7] M. Schrade and T. G. Finstad, *Physica Status Solidi (B)* **1800208**, 1 (2018).

- [8] Z. Chen, X. Zhang, S. Lin, L. Chen, and Y. Pei, *National Science Review*, 888 (2018).
- [9] E. S. Toberer, A. Zevalkink, and G. J. Snyder, *Journal of Materials Chemistry* **21**, 15843 (2011).
- [10] L. Lindsay, D. A. Broido, and T. L. Reinecke, *Phys. Rev. B* **88**, 144306 (2013).
- [11] J. D. Chung, A. J. H. McGaughey, and M. Kaviani, *Journal of Heat Transfer* **126**, 376 (2004), https://asmedigitalcollection.asme.org/heattransfer/article-pdf/126/3/376/5790721/376_1.pdf.
- [12] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. L. Billinge, E. Holm, S. P. Ong, and C. Wolverton, *npj Computational Materials* **8**, 59 (2022).
- [13] T. Xie and J. C. Grossman, *Phys. Rev. Lett.* **120**, 145301 (2018).
- [14] K. Choudhary and B. DeCost, *npj Computational Materials* **7**, 185 (2021).
- [15] V. Fung, J. Zhang, E. Juarez, and B. G. Sumpter, *npj Computational Materials* **7**, 84 (2021).
- [16] C. W. Park and C. Wolverton, *Phys. Rev. Materials* **4**, 063801 (2020).
- [17] G. J. Ackland and A. P. Jones, *Phys. Rev. B* **73**, 054104 (2006).
- [18] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, *The Journal of Chemical Physics* **148**, 241722 (2018), <https://doi.org/10.1063/1.5019779>.
- [19] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, *Chemistry of Materials* **31**, 3564 (2019), <https://doi.org/10.1021/acs.chemmater.9b01294>.
- [20] S. K. Kauwe, J. Graser, A. Vazquez, and T. D. Sparks, *Integrating Materials and Manufacturing Innovation* **7**, 43 (2018).
- [21] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, and O. A. von Lilienfeld, *Journal of Chemical Theory and Computation* **13**, 5255 (2017), pMID: 28926232, <https://doi.org/10.1021/acs.jctc.7b00577>.
- [22] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, in *Proceedings of the 34th International Conference on Machine Learning - Volume 70, ICML'17 (JMLR.org, 2017)* p. 1263–1272.
- [23] F. Legrain, J. Carrete, A. van Roekeghem, S. Curtarolo, and N. Mingo, *Chemistry of Materials* **29**, 6220 (2017), <https://doi.org/10.1021/acs.chemmater.7b00789>.
- [24] X. Qian and R. Yang, *Materials Science and Engineering: R: Reports* **146**, 100642 (2021).

- [25] S. A. Tawfik, O. Isayev, M. J. S. Spencer, and D. A. Winkler, *Advanced Theory and Simulations* **3**, 1900208 (2020), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201900208>.
- [26] R. Jaafreh, Y. S. Kang, and K. Hamad, *ACS Applied Materials & Interfaces* **13**, 57204 (2021).
- [27] S. Ju, R. Yoshida, C. Liu, S. Wu, K. Hongo, T. Tadano, and J. Shiomi, *Phys. Rev. Materials* **5**, 053801 (2021).
- [28] S. Kong, D. Guevarra, C. P. Gomes, and J. M. Gregoire, *Applied Physics Reviews* **8**, 021409 (2021), <https://doi.org/10.1063/5.0047066>.
- [29] S. Kong, F. Ricci, D. Guevarra, J. B. Neaton, C. P. Gomes, and J. M. Gregoire, *Nature Communications* **13**, 10.1038/s41467-022-28543-x (2022).
- [30] M. Gastegger, J. Behler, and P. Marquetand, *Chem. Sci.* **8**, 6924 (2017).
- [31] C. D. Rankine, M. M. M. Madkhali, and T. J. Penfold, *The Journal of Physical Chemistry A*, *The Journal of Physical Chemistry A* **124**, 4263 (2020).
- [32] C. D. Rankine and T. J. Penfold, *The Journal of Chemical Physics* **156**, 164102 (2022).
- [33] K. Ghosh, A. Stuke, M. Todorović, P. B. Jørgensen, M. N. Schmidt, A. Vehtari, and P. Rinke, *Advanced Science* **6**, 1801367 (2019), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.201801367>.
- [34] P. R. Kaundinya, K. Choudhary, and S. R. Kalidindi, *JOM* **74**, 1395 (2022).
- [35] Z. Chen, N. Andrejevic, T. Smidt, Z. Ding, Q. Xu, Y.-T. Chi, Q. T. Nguyen, A. Alatas, J. Kong, and M. Li, *Advanced Science* **8**, 2004214 (2021), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202004214>.
- [36] T. Xie and J. C. Grossman, *Phys. Rev. Lett.* **120**, 145301 (2018).
- [37] ALIGNN GitHub Repository, <https://github.com/usnistgov/alignn>, accessed: 2023-01-23.
- [38] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [39] J. Klimeš, D. R. Bowler, and A. Michaelides, *Journal of Physics: Condensed Matter* **22**, 022201 (2009).
- [40] K. Choudhary and F. Tavazza, *Computational Materials Science* **161**, 300 (2019).
- [41] A. Togo and I. Tanaka, *Scr. Mater.* **108**, 1 (2015).
- [42] M. T. Agne, K. Imasato, S. Anand, K. Lee, S. K. Bux, A. Zevalkink, A. J. Rettie, D. Y. Chung, M. G. Kanatzidis, and G. J. Snyder, *Materials Today Physics* **6**, 83 (2018).

- [43] B. Fultz, *Progress in Materials Science* **55**, 247 (2010).
- [44] S. Guo, S. Anand, Y. Zhang, and G. J. Snyder, *Chemistry of Materials* **32**, 4767 (2020), <https://doi.org/10.1021/acs.chemmater.0c01404>.
- [45] A. Manzoor and D. S. Aidhy, *Materialia* **12**, 100804 (2020).
- [46] X. Hua, S. Hao, and C. Wolverton, *Phys. Rev. Materials* **2**, 095402 (2018).
- [47] S.-i. Tamura, *Physical Review B* **27**, 858 (1983).
- [48] R. Gurunathan, R. Hanus, M. Dylla, A. Katre, and G. J. Snyder, *Phys. Rev. Applied* **13**, 034011 (2020).
- [49] S.-i. Tamura, *Physical Review B* **30**, 849 (1984).
- [50] Note1, See Supplemental Material at [URL will be inserted by publisher] for additional examples of predicted phonon spectra, tabulated values of predicted phonon properties, and additional details about the JARVIS-DFT phonon database.
- [51] D. Bansal, J. L. Niedziela, S. Calder, T. Lanigan-Atkins, R. Rawl, A. H. Said, D. L. Abernathy, A. I. Kolesnikov, H. Zhou, and O. Delaire, *Nature Physics* **16**, 669 (2020).
- [52] R. Ben Sadok, D. Hammoutène, and N. Plugaru, *physica status solidi (b)* **258**, 2000289 (2021), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pssb.202000289>.
- [53] J. Leitner, P. Voňka, D. Sedmidubský, and P. Svoboda, *Thermochimica Acta* **497**, 7 (2010).
- [54] A. T. M. G. Mostafa, J. M. Eakman, M. M. Montoya, and S. L. Yarbrow, *Industrial & Engineering Chemistry Research*, *Industrial & Engineering Chemistry Research* **35**, 343 (1996).
- [55] P. E. Hopkins, *AIP Advances* **1**, 041705 (2011).
- [56] M. Asheghi, W. Liu, and K. E. Goodson, *Proceedings of ICEME ASME International Mechanical Engineering Congress and Exposition*, **113** (2004), https://asmedigitalcollection.asme.org/IMECE/proceedings-pdf/IMECE2004/47071/113/2619615/113_1.pdf.