

Lawrence Berkeley National Laboratory

LBL Publications

Title

Direct Prediction of Phonon Density of States With Euclidean Neural Networks

Permalink

<https://escholarship.org/uc/item/1rn885jc>

Journal

Advanced Science, 8(12)

ISSN

2198-3844

Authors

Chen, Zhantao

Andrejevic, Nina

Smidt, Tess

et al.

Publication Date

2021-06-01

DOI

10.1002/advs.202004214

Peer reviewed

Direct Prediction of Phonon Density of States With Euclidean Neural Networks

Zhantao Chen, Nina Andrejevic, Tess Smidt, Zhiwei Ding, Qian Xu, Yen-Ting Chi, Quynh T. Nguyen, Ahmet Alatas, Jing Kong, and Mingda Li*

Machine learning has demonstrated great power in materials design, discovery, and property prediction. However, despite the success of machine learning in predicting discrete properties, challenges remain for continuous property prediction. The challenge is aggravated in crystalline solids due to crystallographic symmetry considerations and data scarcity. Here, the direct prediction of phonon density-of-states (DOS) is demonstrated using only atomic species and positions as input. Euclidean neural networks are applied, which by construction are equivariant to 3D rotations, translations, and inversion and thereby capture full crystal symmetry, and achieve high-quality prediction using a small training set of $\approx 10^3$ examples with over 64 atom types. The predictive model reproduces key features of experimental data and even generalizes to materials with unseen elements, and is naturally suited to efficiently predict alloy systems without additional computational cost. The potential of the network is demonstrated by predicting a broad number of high phononic specific heat capacity materials. The work indicates an efficient approach to explore materials' phonon structure, and can further enable rapid screening for high-performance thermal storage materials and phonon-mediated superconductors.

One central objective of materials science is to establish structure-property relationships; that is, how specific atomic arrangements lead to certain macroscopic functionalities. This question is historically addressed through trial-and-error of a combination of structure and property characterization, theory, and calculation. However, recent advances in machine learning (ML) suggest a paradigm shift in how structure-property relationships can be directly constructed.^[1,2] To date, ML has seen success in a growing spectrum of materials applications, including materials discovery and design,^[3–6] process automation and optimization,^[7,8] and prediction of materials' mechanical (elastic moduli),^[9–12] thermodynamic and thermal transport (formation enthalpy, thermal conductivity, Debye temperature, heat capacity),^[10,12–16] and electronic (bandgap, superconductivity, topology) properties,^[11,17–24] and atomistic

Z. Chen, N. Andrejevic, Prof. M. Li
Quantum Matter Group
MIT
Cambridge, MA 02139, USA
E-mail: mingda@mit.edu

Z. Chen, Q. Xu
Department of Mechanical Engineering
MIT
Cambridge, MA 02139, USA

N. Andrejevic, Z. Ding, Y.-T. Chi
Department of Materials Science and Engineering
MIT
Cambridge, MA 02139, USA

Dr. T. Smidt
Computational Research Division
Lawrence Berkeley National Laboratory
Berkeley, CA 94720, USA


Dr. T. Smidt
Center for Advanced Mathematics for Energy Research Applications
Lawrence Berkeley National Laboratory
Berkeley, CA 94720, USA

Q. T. Nguyen
Department of Physics
MIT
Cambridge, MA 02139, USA

Dr. A. Alatas
Advanced Photon Source
Argonne National Laboratory
Lemont, IL 60439, USA

Prof. J. Kong
Department of Electrical Engineering and Computer Science
MIT
Cambridge, MA 02139, USA

Prof. M. Li
Department of Nuclear Science and Engineering
MIT
Cambridge, MA 02139, USA

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/advs.202004214>

© 2021 The Authors. Advanced Science published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/advs.202004214

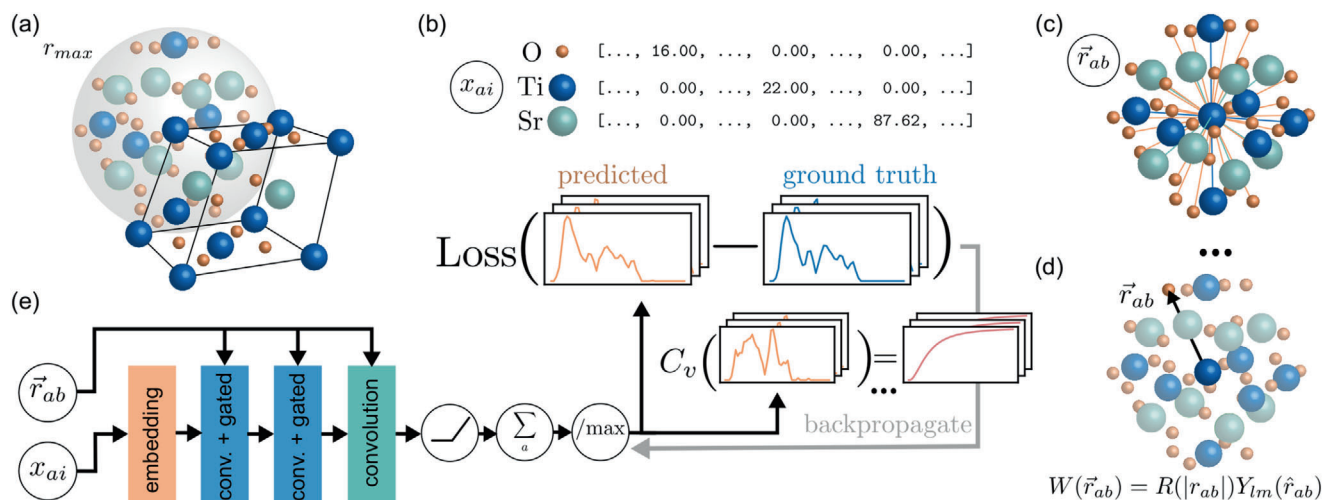


Figure 1. Overview of the E(3)NN architecture for phonon DoS prediction. a) Crystals are converted to periodic graphs by considering all periodic neighbors within a radial cutoff $r_{max} = 5\text{ \AA}$. The example of SrTiO_3 is shown. b) Atom types are encoded as a mass-weighted one-hot encoding. c) Edges join neighboring atoms and store the relative distance vector from the central atom to neighbor. d) The radial distance vectors are used for the continuous convolutional filters $W(\vec{r}_{ab})$ comprising learned radial functions and spherical harmonics. e) The E(3)NN operates on the node and edge features using convolution and gated nonlinear layers. The result is passed to a final activation, aggregation, and normalization to generate the predicted output. The network weights are trained by minimizing the loss function between the predicted and ground-truth phonon DoS.

potentials (potential energy surfaces and force constants).^[25–31] Most property prediction studies consider a low-dimensional output consisting of one or few discrete points. However, the prediction of continuous properties from limited input information remains challenging due to the output complexity and finite data volume. Moreover, for crystalline solids, the crystallographic symmetry poses additional constraints on a generic neural network.

In this work, we build a ML-based predictive model that directly outputs the phonon density of states (DoS) using atomic structures as input. Phonon DoS is a key determinant of materials' specific heat and vibrational entropy and plays a crucial role in interfacial thermal resistance.^[32] It is also tightly linked to thermal and electrical transport^[33] and superconductivity.^[34] However, the acquisition of experimental and computed phonon DoS is nontrivial due to limited inelastic scattering facility resources and high computational cost of ab initio calculations for complex materials.^[33,35] Moreover, the phonon calculations in alloys systems pose significant challenge. Some existing approaches like virtual crystal approximation (VCA) can fail both qualitatively and quantitatively without well-controlled approximations.^[36] This calls for an approach that acquires phonon DoS more efficiently, especially for alloy systems. To build such a model, we employ a Euclidean neural network (E(3)NN) which naturally operates on 3D geometry and is equivariant to 3D translations, rotations, and inversion.^[37–40] E(3)NNs preserve all geometric information of the input and eliminate the need for expensive (approximately 500 fold) data augmentation. Additionally, all crystallographic symmetries of input data are preserved by the network.^[41] In this work, we use E(3)NNs as implemented in the open-source e3nn repository^[40] which merges implementations of Ref. [37] and Ref. [39] and additionally implements inversion symmetry. High-fidelity phonon DoS predictions are achieved using the density functional perturbation theory (DFPT)-based phonon database^[42] containing phonon DoS data of approximately 1,500 crys-

talline solids. Our predictive model can capture the main features of phonon DoS, even for crystalline solids with unseen elements. By predicting the phonon DoS in 4,346 new crystal structures, we identify a list of high heat capacity materials, supported by additional DFPT calculations. Our work offers an efficient technique to acquire phonon DoS directly from atomic structure, making it suitable for high throughput materials design with desirable phonon-related properties.

Crystal structures operated on by the E(3)NN are first converted into a periodic graphs where atoms are nodes N with edges E connecting neighboring atoms within a specified radial cutoff, including periodic images (Figure 1). Each edge $e_{ab} \in E$ stores the radial distance vector between atom a and neighbor b , \vec{r}_{ab} , up to some radial cutoff $|r_{max}|$, and is used by the convolutional kernels of the E(3)NN. The input node features are scalars that captures its atomic type and mass using one-hot encoding; for instance, a hydrogen atom is encoded as $x_{\text{H}} = [m_{\text{H}}, 0, \dots, 0]$. After an initial embedding layer which takes the 118-length one-hot mass-weighted encodings to 64 scalar features, the constructed graph is then passed to the E(3)NN, which iteratively operates on the features with multiple “Convolution and Gated Block” layers as described for the L1Net of ref. [43] (see Supporting Information for more details). After the final layer, which consists of only a convolution, all resulting node features are summed and passed through a final activation (ReLU) and normalization layer to predict the phonon DoS, comprising 51 scalars. The absolute magnitude of the phonon DoS can easily be recovered from the normalized DoS by noticing that $\int g(\omega)d\omega = 3N$, where N is the number of atoms in the unit cell; thus, we ensure that normalization of the DoS does not compromise meaningful prediction. The E(3)NN weights are optimized by minimizing the mean squared error (MSE) loss function between the DFPT-computed DoS g and E(3)NN-predicted \hat{g} . The full network structure is provided in the Supporting Information.

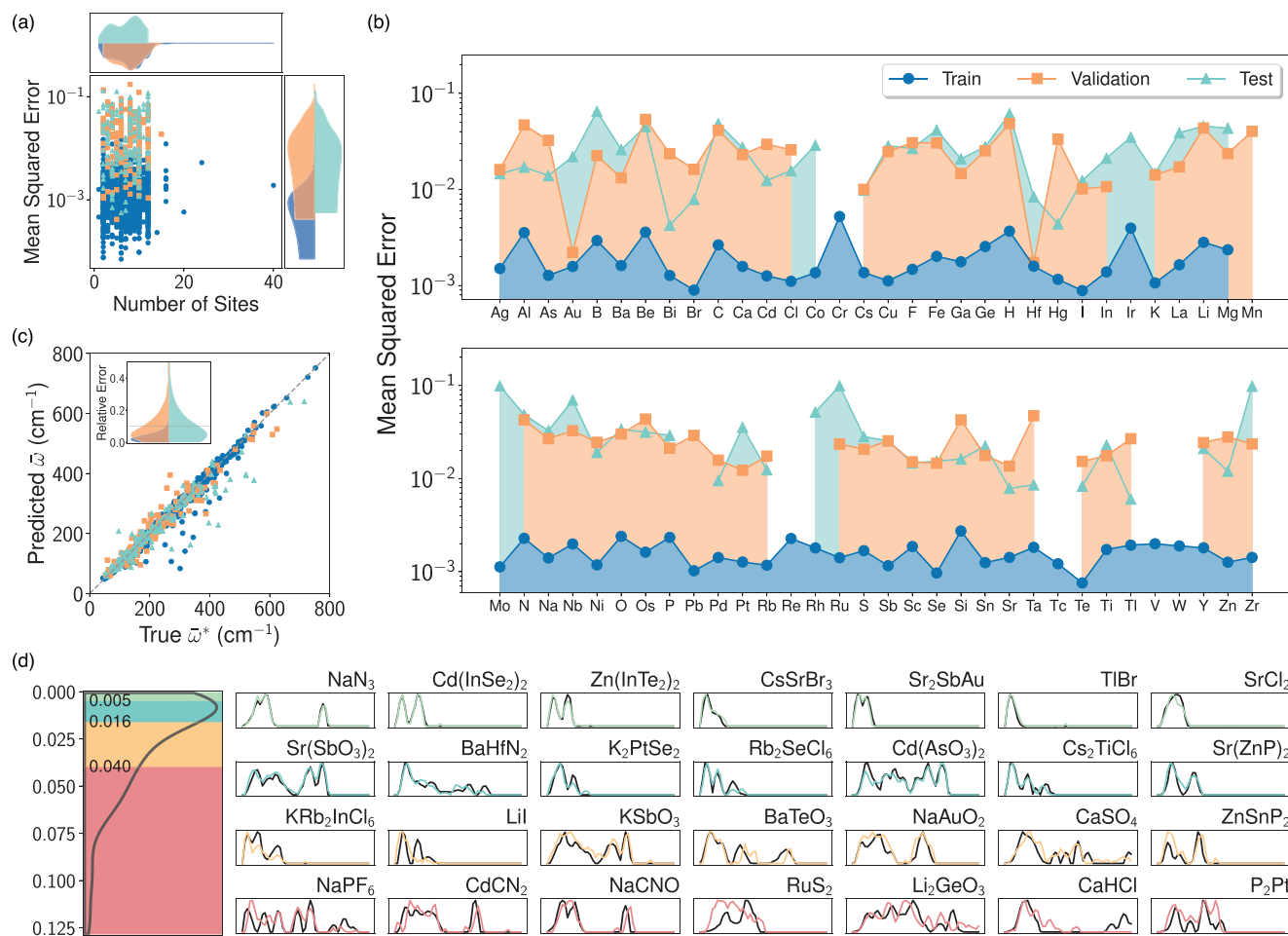


Figure 2. Performance of the Euclidean neural network-based predictive model. a) Mean squared error versus total number of sites in a unit cell in training (blue), validation (orange), and test (green) sets. b) Average mean squared error of compounds containing each element. c) Comparison between E(3)NN-predicted average phonon frequency and ground truth. The inset shows the relative error $|\bar{\omega} - \bar{\omega}^*|/\bar{\omega}^*$ distribution of the three datasets. d) Randomly selected examples in the test set within each error quartile. (Left) MSE distribution showing that it is heavily peaked in the 1st and 2nd quartiles with lower error.

We perform several analyses to evaluate our model given the limited training data. **Figure 2a** shows that there is no obvious correlation between the MSE and the number of basis atoms within unit cells among training, validation, and test datasets (additional statistics in are available in the Supporting Information). The overall test set error is higher compared to the training set but similar to that of the validation set, suggesting good generalizability. We also present the MSE as a function of different elements (**Figure 2b**) and observe comparable error levels, indicating balanced prediction. Lastly, we compute the average phonon frequency $\bar{\omega} = \frac{\int d\omega g(\omega)\omega}{\int d\omega g(\omega)}$ for both E(3)NN-predicted and DFPT ground-truth spectra, which show excellent agreement on the test set (**Figure 2c**); specifically, for 70% of the testing samples, the relative error is below 10%. This strongly suggests the capability of our model to predict phonon DoS.

To visualize the model performance, we plot seven randomly selected examples from the test set in each error quartile in **Figure 2d**, with rows 1 through 4 corresponding to the 1st quartile with highest agreement through the 4th quartile with lowest

agreement, respectively. Additional examples are plotted in the Supporting Information. The predicted DoS in the 1st and 2nd quartiles show excellent agreement with DFPT calculations by reproducing fine features, while the 3rd and 4th quartiles show good or acceptable agreement by capturing main features. For instance, the predicted DoS of NaPF₆, CdCN₂, and NaCNO capture the energy of acoustic and optical phonon branches well but mispredict the relative amplitudes of certain peaks. Nonetheless, the phonon bandgap, a key quantity to determine phonon-phonon scattering, can be accurately extracted. Similarly, for KSbO₃ and Li₂GeO₃, the predictions exhibit broadband DoS distribution, agreeing with DFPT calculations. A large discrepancy can be seen for RuS₂, yet the bandwidth agreement is still good although the 0.099 MSE is among the largest errors in the test set (**Figure 2a**). The good test set performance and generalizability suggest the suitability of our model to predict phonon DoS for a broad range of new materials.

We compare the E(3)NN predictions in six materials with experimental DoS data available from inelastic scattering (**Figure 3**). Given the disorder and anharmonic effects in a measured

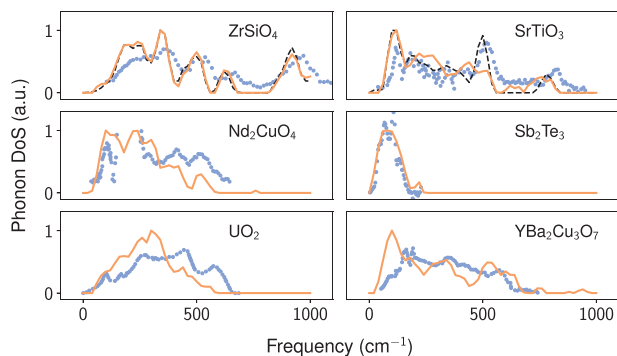


Figure 3. Comparison between E(3)NN model predictions (orange curves) and inelastic scattering DoS data (blue dots). ZrSiO₄ was in the training set and SrTiO₃ in the test set (black dashed lines denote corresponding DFPT results.^[42]) The remaining examples were absent in all datasets used for training, validation, and testing, and contain two unseen elements, Nd and U. Reproduced with permission.^[44–46,48,49] Copyrights 2008, 1991, 1993, 2014, 2016, American Physical Society;^[47] Copyright 1981, IOP Publishing;^[50] Copyright 1997, Elsevier B.V.

sample, disagreement between DFPT calculations and measured data can happen. As a result, lower agreement is expected between experimental and E(3)NN-predicted DoS since the ground-truths are based on DFPT calculations.

Although the E(3)NN-predicted DoS do not match the fine features of the experimental spectra, several key features (peak positions, gaps, and energy bandwidths) are still well-predicted and can be valuable in guiding experimental planning, which can serve as useful guidance for planning inelastic neutron and x-ray scattering measurements, where experimental resources are largely limited to national laboratory facilities.

One of the most important applications of our prediction model may lie in the predictive power in alloy systems, particularly crystalline alloys with substitutional disorders. For example, given a binary alloy with composition A_pB_{1-p} (0 ≤ p ≤ 1), the input alloy encoding vector can take the following form

$$x_{\text{alloy}} = [0, \dots, pm_A, \dots, (1-p)m_B, \dots, 0] \quad (1)$$

where the two-hot encoding pm_A and $(1-p)m_B$ are located at the vector indices corresponding the atomic numbers of A and B, respectively, weighted by composition. With this definition of Eq. 1, it can be directly reduced to pure phase one-hot encoding A (or B) by simply setting $p = 1$ (or $p = 0$), and it can be generalized to more complicated alloys directly. In fact, Equation (1) contains the essence of VCA with both mass average $m_{\text{VCA}} = pm_A + (1-p)m_B$ and scattering potential average $V_{\text{VCA}} = pV_A + (1-p)V_B$, but combined into one equation: the mass effect is contained in the numerical values of pm_A and $(1-p)m_B$, and the potential effect is encoded by turning on the vector indices that correspond to the atomic species A and B. Since the computation of pure phase and alloy differs only by atomic embedding method, the alloy calculation does not generate any additional computational cost.

We demonstrate the power of this approach with the alloy Mg₃Sb_{2(1-p)}Bi_{2p} with $p \in [0, 1]$. The model evaluation is done by the aforementioned two-hot encoding for Sb on top of the structure Mg₃Bi₂, with simultaneously interpolating the lattice con-

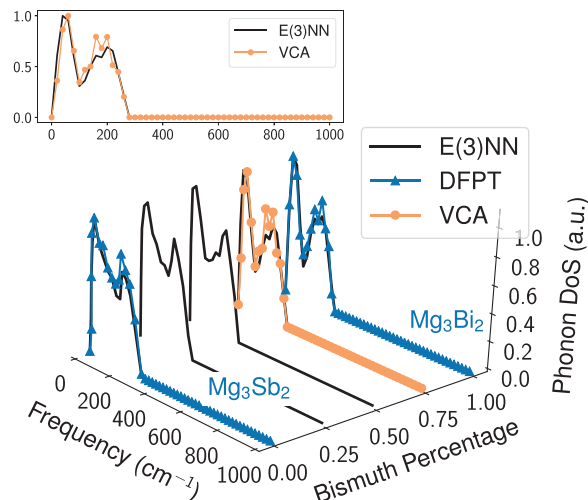


Figure 4. Comparison between E(3)NN model predictions and virtual crystal approximation (VCA) calculations. E(3)NN model can predict phonon DoS for the alloy Mg₃Sb_{2(1-p)}Bi_{2p} of continuous p with two-hot weighted encoding. The triangle-marked curves indicate DFPT results for Mg₃Sb₂ and Mg₃Bi₂. Reproduced with permission.^[42] Copyright 2018, Springer Nature;^[51] Copyright 2018, Elsevier B.V. The circle-marked curve represent VCA calculations for Mg₃Sb_{0.5}Bi_{1.5} ($p = 0.75$). The inset figure shows the front view for E(3)NN and VCA comparison.

starts in-between values of two limit structures (namely, Mg₃Bi₂ for $p = 1$ and Mg₃Sb₂ for $p = 0$) according to the composition. In this case, both input vectors and the structure are recovered to the pure phase Mg₃Sb₂ when $p = 0$ and vice versa. We compute the phonon DoS in alloy Mg₃Sb_{0.5}Bi_{1.5} and compare it with VCA calculations, as shown in Figure 4, where the peak positions and magnitudes are both well predicted. The E(3)NN model used to evaluate this alloy system is trained with an additional Mg₃Bi₂ phonon DoS curated from^[51] compared with the model used in the rest of this paper, while Mg₃Sb₂ has already been included in the original training set.

We apply the predictive model on 4346 unseen crystal structures without ground-truth DoS from the Materials Project.^[52] The data consistency check is performed (Supporting Information), showing a reasonable trend of an elastic spring model. Figure 5a illustrates the average phononic specific heat capacity C_V of crystalline solids containing a given element, using the relation^[53]

$$C_V(T) = \frac{k_B}{m_{\text{tot}}} \int_0^\infty \left(\frac{\hbar\omega}{2k_B T} \right)^2 \text{csch}^2 \left(\frac{\hbar\omega}{2k_B T} \right) g(\omega) d\omega \quad (2)$$

where m_{tot} is the total mass of all N atoms in the unit cell, and the phonon DoS is normalized such that $\int g(\omega) d\omega = 3N$. Materials containing light elements tend to have high heat capacity, which is reasonable. The distribution of C_V evaluated from Equation (2) is shown in Figure 5b, where $\approx 2\%$ of materials show a C_V greater than 1000 J/(kgK)⁻¹. The inset shows the average phonon DoS of highest- C_V materials. Materials with higher C_V appear to have high spectral weight at higher energies, consistent with expectation. This trend is also noticed by inspecting the scatter plot of phonon DoS along the first two principal components (Figure 5c), where high heat capacity materials appear clustered with

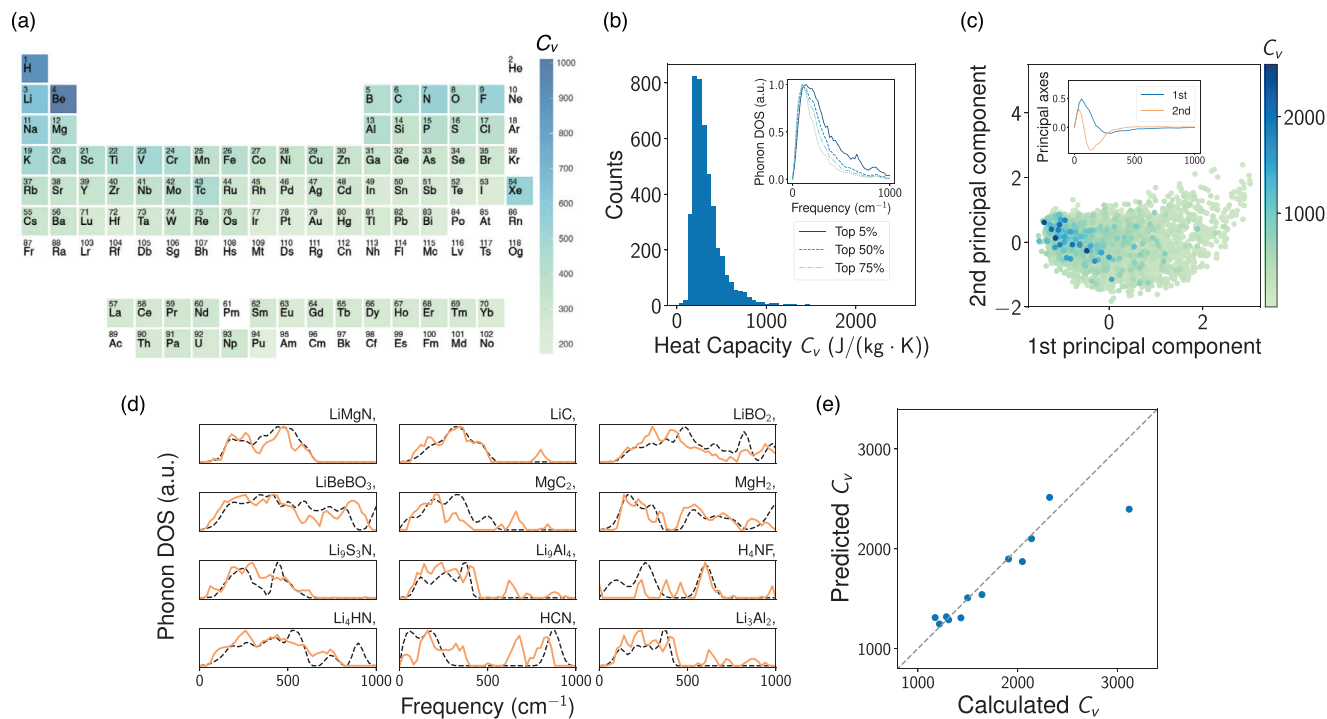


Figure 5. The search for high specific heat capacity (C_V) materials. a) Periodic table colored by average C_V of materials containing each element. b) Histogram showing distribution of C_V evaluated from E(3)NN-predicted phonon DoS. The inset illustrates average phonon DoS of materials with highest C_V . c) Distribution of predicted phonon DoS along the first two principal components, colored by the E(3)NN-predicted C_V magnitudes. The inset shows first two principal axes in the original frequency basis. d) Comparison between E(3)NN-predicted and DFPT-computed phonon DoS, the dashed black curves represent DFPT results. e) 2D histogram comparing specific heat capacities evaluated from E(3)NN-predicted and DFPT-calculated phonon DoS.

respect to the first principal axis. The first principal axis has a broad negative peak extending to high energies; thus, the clustering of high C_V materials in the negative first principal direction parallels the shift of their phonon DoS toward higher energies.

To validate our model's predictions of high C_V materials, we select 12 materials with ultrahigh predicted C_V and carry out independent DFPT calculations. Since the maximum frequency of the training data was set to 1000 cm^{-1} , which is sufficient for the majority of materials, the C_V evaluated by DFPT was also cut off at 1000 cm^{-1} for fair comparison. The DoS comparisons between E(3)NN and DFPT in these high- C_V materials are shown in 5d, where satisfactory agreements are achieved in most examples except for H_4NF and HCN . The C_V at room temperature $T = 293.15\text{ K}$ evaluated from E(3)NN predictions and DFPT calculations are plotted in Figure 5e, showing excellent agreement for most materials. Values of the plotted C_V are presented in Table S2, Supporting Information, together with C_V computed from full energy ranges for comparison. We attribute the large discrepancy in hydrogen- and lithium-rich materials to the electrostatic effect of hydrogen and lithium bondings beyond mass effect,^[54] while the current model mainly considers the mass effect.

In this work, we present a machine learning-based predictive model to directly acquire the high-dimensional material property of phonon DoS, using only “first-principles” inputs, namely atomic species and positions. Due to their equivariance, Euclidean neural networks are able to capture the symmetries of the input crystal, making them data-efficient. A small training set of only 1200 examples is sufficient to generate meaningful predic-

tions, outperforming a well-trained convolutional neural network even with data augmentation (Supporting Information). One limitation of the current model is identified as samples with elastic strain, potentially related to the fact that we confined output energy range to positive values (Supporting Information, which also includes Refs [56–58]). Even so, Euclidean neural networks can be applied to predicting broader properties in crystalline solids, where there are often issues of data scarcity. Most importantly, our atomic embedding approach offers an extremely efficient way to compute phonon DoS in alloys, where the alloy prediction has the same computational cost as pure phase by simply changing the 1-hot embedding to a weighted multi-hot embedding.

In contrast to ab initio calculations and inelastic scattering that acquire phonon DoS deterministically, a ML model is data-driven and probabilistic in nature. It is thus impractical to fully rely on a ML-based predictive model to acquire materials properties without further validation. However, the power of the ML approach goes far beyond obtaining property magnitudes at the individual material level. From a materials design perspective, ML demonstrates extremely high efficiency in rapidly screening candidates with a target property. In our case, the prediction of phonon DoS on the 4346 unseen materials can be done in less than 30 minutes on a single entry-level GPU. From a property optimization perspective, instead of measuring an individual material with high precision, the ML approach searches outputs high-performance candidates in a batch, offering more choices. From an experimental perspective, the ML model can be valuable in guiding the experimental planning with limited national facility resources.

From an application point-of-view, the Dulong-Petit law poses a grand challenge in searching for promising materials for thermal storage,^[55] and a highly efficient approach can further support an inverse design from property to desirable structures. In sum, our model provides a promising framework to enable high-throughput screening and guide experimental planning for materials with exceptional thermal properties. It further sheds light on elucidating the fundamental links between symmetry, structure, and elementary excitations in condensed matter.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Z.C., N.A., and T.S. contributed equally to this work. Z.C., N.A., and M.L. thank C.H. Rycroft and J.W. Lynn for helpful discussions. Z.C., N.A., and M.L. acknowledge the support from U.S. DOE, Office of Science (SC), Basic Energy Sciences (BES), award No. DE-SC0020148. N.A. acknowledges the support of the National Science Foundation Graduate Research Fellowship Program under Grant No. 1122374. T.S. acknowledges support from the Laboratory Directed Research and Development Program and the Center for Advanced Mathematics for Energy Research Applications of Lawrence Berkeley National Laboratory both under U.S. Department of Energy Contract No. DE-AC02-05CH11231. Y.-T.C. thanks K. Van Vliet for computational support.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in [GitHub] at [https://github.com/zhantaochen/phonondos_e3nn].

Keywords

density of states, machine learning, phonons

Received: November 25, 2020

Revised: February 1, 2021

Published online: March 16, 2021

- [1] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, *npj Comput. Mater.* **2017**, *3*, 1.
- [2] J. Schmidt, M. R. Marques, S. Botti, M. A. Marques, *npj Comput. Mater.* **2019**, *5*, 1.
- [3] P. Raccuglia, K. C. Elbert, P. D. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, *Nature* **2016**, *533*, 73.
- [4] A. O. Oliynyk, E. Antono, T. D. Sparks, L. Ghadbeigi, M. W. Gaultois, B. Meredig, A. Mar, *Chem. Mater.* **2016**, *28*, 7324.
- [5] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, A. Aspuru-Guzik, *ACS Cent. Sci.* **2018**, *4*, 268.
- [6] Y. Liu, T. Zhao, W. Ju, S. Shi, *J. Materiomics* **2017**, *3*, 159.
- [7] F. Häse, L. M. Roch, A. Aspuru-Guzik, *Chem. Sci.* **2018**, *9*, 7642.
- [8] J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, *Nature* **2018**, *559*, 377.
- [9] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, R. Ramprasad, *Sci. Rep.* **2013**, *3*, 1.
- [10] O. Isayev, C. Oses, C. Toher, E. Gossett, S. Curtarolo, A. Tropsha, *Nat. Commun.* **2017**, *8*, 1.
- [11] T. Xie, J. C. Grossman, *Phys. Rev. Lett.* **2018**, *120*, 145301.
- [12] C. Chen, W. Ye, Y. Zuo, C. Zheng, S. P. Ong, *Chem. Mater.* **2019**, *31*, 3564.
- [13] J. Carrete, W. Li, N. Mingo, S. Wang, S. Curtarolo, *Phys. Rev. X* **2014**, *4*, 011019.
- [14] A. van Roekeghem, J. Carrete, C. Oses, S. Curtarolo, N. Mingo, *Phys. Rev. X* **2016**, *6*, 041061.
- [15] S. A. Tawfik, O. Isayev, M. J. Spencer, D. A. Winkler, *Adv. Theory Simul.* **2020**, *3*, 1900208.
- [16] B. Mortazavi, E. V. Podryabinkin, S. Roche, T. Rabczuk, X. Zhuang, A. V. Shapeev, *Mater. Horiz.* **2020**.
- [17] L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, *npj Comput. Mater.* **2016**, *2*, 16028.
- [18] Y. Zhuo, A. Mansouri Tehrani, J. Brgoch, *J. Phys. Chem. Lett.* **2018**, *9*, 1668.
- [19] Y. Dong, C. Wu, C. Zhang, Y. Liu, J. Cheng, J. Lin, *npj Comput. Mater.* **2019**, *5*, 1.
- [20] B. Meredig, E. Antono, C. Church, M. Hutchinson, J. Ling, S. Paradiso, B. Blaiszik, I. Foster, B. Gibbons, J. Hattract-Simpers, A. Mehta, L. Ward, *Mol. Syst. Des. Eng.* **2018**, *3*, 819.
- [21] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, I. Takeuchi, *npj Computational Materials* **2018**, *4*, 29.
- [22] N. Andrejevic, J. Andrejevic, C. H. Rycroft, M. Li, arXiv preprint, arXiv:2003.00994 **2020**.
- [23] N. Claussen, B. A. Bernevig, N. Regnault, *Phys. Rev. B* **2020**, *101*, 245117.
- [24] M. S. Scheurer, R.-J. Slager, *Phys. Rev. Lett.* **2020**, *124*, 226401.
- [25] Z. Li, J. R. Kermode, A. De Vita, *Phys. Rev. Lett.* **2015**, *114*, 096405.
- [26] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, K.-R. Müller, *Sci. Adv.* **2017**, *3*, e1603015.
- [27] I. Kruglov, O. Sergeev, A. Yanilkin, A. R. Oganov, *Sci. Rep.* **2017**, *7*, 1.
- [28] A. Glielmo, P. Sollich, A. De Vita, *Phys. Rev. B* **2017**, *95*, 214302.
- [29] V. Botu, R. Batra, J. Chapman, R. Ramprasad, *J. Phys. Chem. C* **2017**, *121*, 511.
- [30] F. Legrain, A. van Roekeghem, S. Curtarolo, J. Carrete, G. K. Madsen, N. Mingo, *J. Chem. Inf. Model.* **2018**, *58*, 2460.
- [31] L. Zhang, D.-Y. Lin, H. Wang, R. Car, E. Weinan, *Phys. Rev. Mater.* **2019**, *3*, 023804.
- [32] Y.-J. Wu, L. Fang, Y. Xu, *npj Comput. Mater.* **2019**, *5*, 56.
- [33] F. Giustino, *Rev. Mod. Phys.* **2017**, *89*, 015003.
- [34] J. Bardeen, L. N. Cooper, J. R. Schrieffer, *Phys. Rev.* **1957**, *108*, 1175.
- [35] S. Baroni, S. De Gironcoli, A. Dal Corso, P. Giannozzi, *Rev. Mod. Phys.* **2001**, *73*, 515.
- [36] H. R. Seyf, L. Yates, T. L. Bougher, S. Graham, B. A. Cola, T. Detchprohm, M.-H. Ji, J. Kim, R. Dupuis, W. Lv, A. Henry, *npj Comput. Mater.* **2017**, *3*, 49.
- [37] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, P. Riley, arXiv e-prints, arXiv:1802.08219 **2018**.
- [38] R. Kondor, Z. Lin, S. Trivedi, in *Advances in Neural Information Processing Systems*, vol. 32. ACM, New York **2018**, pp. 10117–10126, <https://dl.acm.org/doi/proceedings/10.5555/3326943>.
- [39] M. Weiler, M. Geiger, M. Welling, W. Boomsma, T. Cohen, in *Advances in Neural Information Processing Systems*, vol. 32. ACM, New York **2018**, pp. 10402–10413, <https://dl.acm.org/doi/proceedings/10.5555/3326943>.

- [40] M. Geiger, T. Smidt, B. K. Miller, W. Boomsma, K. Lapchevskyi, M. Weiler, M. Tyszkiewicz, B. Dice, J. Frelsen, S. Sanborn, M. Alby, *e3nn*: A modular framework for euclidean neural networks, github.com/e3nn/e3nn (accessed: January 2021).
- [41] T. E. Smidt, M. Geiger, B. K. Miller, arXiv e-prints, arXiv:2007.02005 **2020**.
- [42] G. Petretto, S. Dwaraknath, H. P. Miranda, D. Winston, M. Giantomassi, M. J. Van Setten, X. Gonze, K. A. Persson, G. Hautier, G.-M. Rignanese, *Sci. Data* **2018**, *5*, 180065.
- [43] B. K. Miller, M. Geiger, T. E. Smidt, F. Noé, arXiv e-prints, arXiv:2008.08461 **2020**.
- [44] N. Choudhury, E. J. Walter, A. I. Kolesnikov, C.-K. Loong, *Phys. Rev. B* **2008**, *77*, 134111.
- [45] J. Lynn, I. Sumarlin, D. Neumann, J. Rush, J. Peng, Z. Li, *Phys. Rev. Lett.* **1991**, *66*, 919.
- [46] I. Sumarlin, J. Lynn, D. Neumann, J. Rush, C. Loong, J. Peng, Z. Li, *Phys. Rev. B* **1993**, *48*, 473.
- [47] H. Rauh, R. Geick, H. Kohler, N. Nucker, N. Lehner, *J. Phys. C: Solid State Phys.* **1981**, *14*, 2705.
- [48] J. W. Pang, A. Chernatynskiy, B. C. Larson, W. J. Buyers, D. L. Abernathy, K. J. McClellan, S. R. Phillpot, *Phys. Rev. B* **2014**, *89*, 115132.
- [49] S. Chaplot, L. Pintschovius, N. Choudhury, R. Mittal, *Phys. Rev. B* **2006**, *73*, 094308.
- [50] J. Nipko, C.-K. Loong, *Phys. B: Condens. Matter* **1997**, *241*, 415.
- [51] M. T. Agne, K. Imasato, S. Anand, K. Lee, S. K. Bux, A. Zevalkink, A. J. Rettie, D. Y. Chung, M. G. Kanatzidis, G. J. Snyder, *Mater. Today Phys.* **2018**, *6*, 83.
- [52] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
- [53] C. Lee, X. Gonze, *Phys. Rev. B* **1995**, *51*, 8610.
- [54] A. Shahi, E. Arunan, *Phys. Chem. Chem. Phys.* **2014**, *16*, 22935.
- [55] A. Henry, R. Prasher, A. Majumdar, *Nat. Energy* **2020**, *5*, 635.
- [56] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, G. Ceder, *Comput. Mater. Sci.* **2013**, *68*, 314.
- [57] J. Lynn, H. Smith, R. Nicklow, *Phys. Rev. B* **1973**, *8*, 3493.
- [58] J. Hoffmann, L. Maestrati, Y. Sawada, J. Tang, J. M. Sellier, Y. Bengio, arXiv preprint, arXiv:1909.00949 **2019**.