Materials Today Physics 28 (2022) 100868

Contents lists available at ScienceDirect

Materials Today Physics

journal homepage: www.journals.elsevier.com/materials-today-physics

Leveraging low-fidelity data to improve machine learning of sparse high-fidelity thermal conductivity data via transfer learning

Z. Liu [a,b], M. Jiang [c], T. Luo [a,d,*]

[a] Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, 46556, United States
[b] Department of Applied Physics, School of Physics and Electronics, Hunan University, Changsha, 410082, China
[c] Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN, 46556, United States
[d] Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN, 46556, United States

ARTICLE INFO

Keywords:
Thermal conductivity
Machine learning
Transfer learning
First-principles simulation

ABSTRACT

Lattice thermal conductivity (TC) of semiconductors is crucial for various applications, ranging from microelectronics to thermoelectrics. Data-driven approach can potentially establish the critical composition-property relationship needed for fast screening of candidates with desirable TC, but the small number of available data remains the main challenge. TC can be efficiently calculated using empirical models, but they have inferior accuracy compared to the more resource-demanding first-principles calculations. Here, we demonstrate the use of transfer learning (TL) to improve the machine learning models trained on small but high-fidelity TC data from experiments and first-principles calculations, by leveraging a large but low-fidelity data generated from empirical TC models, where the training on high- and low-fidelity TC data are treated as different but related tasks. TL improves the model accuracy by as much as 23% in R^2 and reduces the average factor of difference by as much as 30%. Using the TL model, a large semiconductor database is screened, and several candidates with room temperature TC > 350 W/mK are identified and further verified using first-principles simulations. This study demonstrates that TL can leverage big low-fidelity data as a proxy task to improve models for the target task with high-fidelity but small data. Such a capability of TL may have important implications to materials informatics in general.

1. Introduction

Lattice thermal conductivity (TC) is a critical property of semiconductors, impacting a wide range of applications, such as electronics thermal management [1,2], thermoelectrics [3,4], and thermal barrier coatings [5]. Low TC, for example, is preferred for thermoelectric energy conversion (the figure-of-merit is inversely proportional to TC), and high TC is desirable for electronic cooling to improve device reliability and performance. Identifying semiconductors with desirable TC has mainly relied on physics-guided search, which is usually time-consuming. Recent successes include discovering low [6–9] and high TC [10–14] materials for physics-guided studies [illegible]

materials at state-of-the-art is the first-principles lattice dynamics combined with the phonon Boltzmann Transport Equation (BTE), which can be highly accurate [14–18]. These calculations, however, take a significant amount of time and computational resources, as parameters (e.g., pseudopotential, selection of force constant cutoff, q-point convergence, higher-order phonons, phonon renormalization) need to be carefully tested to obtain reliable results. Although first-principles calculations offer high-fidelity results, the number of semiconductors with TC calculated using this method is still limited. Empirical models have been leveraged for fast estimation of TC for a large number of semiconductors [19]. One example is the Slack model [20], where Debye temperature, Grüneisen parameter, average atomic mass, number of atoms [illegible] are input parameters. Recently, Toher et al. developed Automatic GIBBS Library (AGL) which calculated the Debye temperature and Grüneisen parameter from first-principles, enabling high-throughput calculations of TC using the Slack model [21]. The [illegible]

* Corresponding author. Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, 46556, United States.
E-mail address: tluo@nd.edu (T. Luo).

https://doi.org/10.1016/j.mtphys.2022.100868
Received 23 June 2022; Received in revised form 26 August 2022; Accepted 25 September 2022
Available online 4 October 2022

relatliavrdegaletyaosfeetmiconldautcTtCowciet1h56m7aterialplhso,npomropevritayucahTLstratletwgaysnotetdhatthe
butthefdeloifstycphrediTcCtiesrdelatliowweiltyahpparenrfrst-preinhecicpbharnsdolgdaptwaerienaccudruatetoteehe
underestoimnatvaeiaodnu[21]es Howevseirn,tcheAGL TC is well-kdhoenowsnfiutnyctitibnoeeadorlaynd guanpderestiismsauteior
calcuulsaitaepgphysmnodoeillst,houclodhtsaoimmeeaningf[u35]l HowevtehrfIe,rst-prdiancipcllfaeotsprihomspornoperties
compositionep-atrotpoeenrstlyhips. havbeeepnrovtebnavheigh-fdaelrtdhitetrya,iTnLmeoddelurned
Data-dmraicvheinlnearn(MLmg)ethopdosvaindoetrhoeurteouttobehighalctycurTahteer.eTflocraahne,apotentecicanhhtique
topredmiacttenTCs[21–25] Pionwerrkisncluadses Bayesian evebaigeow-fddealttiatbteylipmprMLvmeodetlrsaiwnietdh
optimizeatbiaoshenTd01 TCdatfaofmrst-prsinmcuiplasensablulhtigh-fdaetlai.ty
tionfseurltraTlCocwompoubndySsekoetal[26] anda In thwiosrwkd,edemonsthraabtieltoutsyethTeLs strategy
high-throsueaghrophuowTCinhalf-Hseeunsilecorndbuyctiompsrotvhe MLmodeal ccurtaracyiwnietdh170 high-fdTeClaitay
Carreetatl.[27]Cheenatll.atuetrileixzpeedrimaenvatalTaClby levera156l7ogw-fdAeGILTiCaysthperockaytTah.e17T0C
datfao1r00semicondutcrtaMiLnmsodealnsdecurfseiavteurdeatcaombienxperimenetasluvraelbuaensdthofseofmrst-
eliminiastppdtriofendlthree ledveasndcr[28]tHoorwsevtehre, princliaptltieysneamciaclscullfartoimobheisterTahteusrveo.
traindaitnigtheps reisotrudairese latsimvaeallyt,howuigthhseodfslaataardeiffeburentlaatstehdesyhoubldotchontianifnor
highdelTihtsym.adlaltcaheatdobiasseodveriatgheehemicanlatioofthnhsetructurer-eplraotpiedrntsiehsig1p5.6A7GL TC
spaaecedhumsayresterxitcrtapochaptaboilfhiterya iMnLed datade,enpeutnreatlw(oDrNkNi)sfrsttaiTnLeidsthe applitoed
modeForsinstamnoeewreoen l8ymateroiuadlfs10 witThC transthfekrnowleodhgeeomposition/strruecltaure
ove100W/mKinChesnwor[k28] potentiimalittlymeogd'el tionsheaipfnreodmhAeGLdattoatherainusingthgle170 high-
abiltiostcyrheingTChmaterials. fdeldiatty(Fig. 1).TLishowtomimprothveDNNmodealccuracy
Iwtouldbeideiaflneahevebagthsemablulhtigh-fdeflrioRmy² = 0.667ashigahs0.833dedutcheasverfagetor
datanbdibgultow-fddealttiatyaMiLmnmodetloisncreehaseediffe(rAeFnDc)eo2m.1t8aslowas1.5U2s.inthgTeLmodeall,l
prediactcionmanodyheangeapplicaTbriahnsitfeya.rnisnegmicondutchMteatesPraoljseactabl[3a6s]ersecreeamnedd,
(TLi)sMLschetmoeevercsommadlaltparobl[e2m9s],andithas sevecraanldiwdaitreosotmemperTaCt>u3r5e0W/mKariede-nti
beesnuccessapfpuliilehdfemwaterianflosrmsatuidTiCelse. fedw,hiahreurtvheerriufsiedfnrgst-prdiancipll@eutsrions
veralkgnoewleodgseeinalargdeatasfaetpropxryopeartnydy resudlemtnsonstthrfatlectahevebraiguelw-fddealtiatpyroxy
thetnransifteoirmsprotvhe MLmodeplerformeantcaergettoimprmooveddelfosrhteargaestkwiththhigh-fbluesltmidatdylata.
propewrittahymuchsmaldlaetra[s3e1t–s34] Imaprevsioduyslyu,cahcapaboifTLimtayyhaviemporitmapnltictaotnioers ials
Lietalu.se~d1400electbraonmdgdaptfarofmrst-prinicniofpolrensiagteincesral.
calcultaotireaaminseumaeltwoarnkdl,tswseigwhietssues eads
inipuaitaalmeitterrasianriontgrmheeurnaeltwwoorrkl[a3b4]phonon 2. Results
bandgdaalpp[32]. Itwasprovtehmatahaeccumraecaysubryeeldan
absoeluterroofpsredicdoulbbeedreduby6c5e%fotrhetarget For101semicondTuCctdoarsa,avaiiliatbbleottehhAeGL

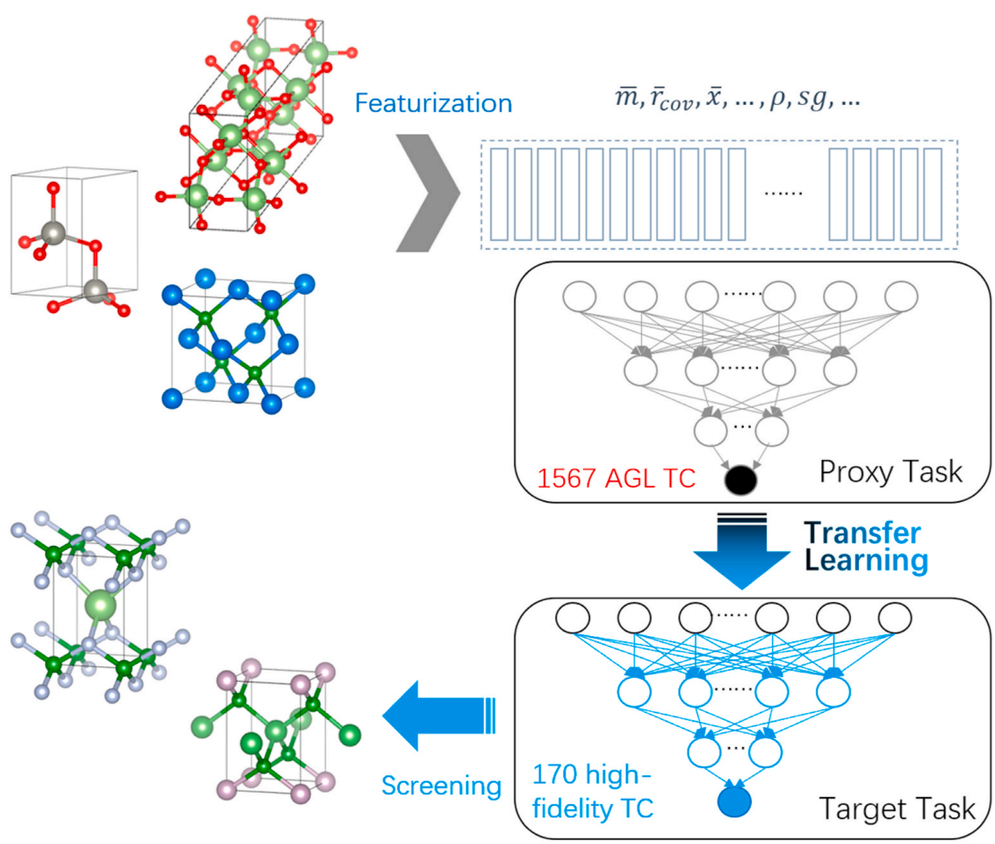

Fig. 1. Workfow of TLScheme of this TLs work fouswith the big low-fidAeGlLiTCaysthe proxy taskand transfer learning model to con structDNN model for the target task with high-fidelity data.

data and the high-fidelity data. We compare the TC values as a parity plot in Fig. 2a. The first observation is that there is a general positive correlation between the AGL data and the high-fidelity data, i.e., semiconductors with higher TC in reality generally have higher TC in the AGL data. This suggests that the AGL TC data has encoded meaningful composition-property relationship, which may be learned by ML. However, AGL underestimates the TC, especially for those with TC > 100 W/mK. On the other hand, the high-fidelity data spreads over four orders of magnitude even with small data quantity, indicating a proper data spread (Fig. 2b). In comparison, the AGL data has a much narrower TC distribution centered around a low mean value (~3 W/mK) as shown in Fig. 2c. However, we examine the representations of the material compositions in the data set with a t-SNE plot (Fig. 2d), which visualizes the distribution of the high-dimensional data in a low-dimensional manifold [37]. It can be seen that the materials with high-fidelity TC only cover a small fraction of the space of the AGL data. Thus, building ML model based on the small high-fidelity data will likely introduce large sampling bias. As a result, we believe that the AGL data set may help achieve ML models that are more generally applicable.

ML model for proxy task: semiconductor compositions are represented as a 65-dimensional vector which concatenates the element composition and descriptors (see Methods). We first pre-train DNNs with the 5,576 AGL TC data with three hidden layers. The number of neurons in the layers of DNNs are randomized (see Methods). A total of 91 shot models are separately trained with the AGL TC data using PyTorch [38]. The model regression performance metrics are evaluated via five-fold cross-validation, and the predictions on validation sets are then compared to the ground truth. For each model, five-fold cross-validation is performed with different training-validation data partitions. The total 5 values of performance metric are then averaged, and the standard deviation is calculated to account for error bar. It should be noted that considering the large orders of magnitude of TC and the relatively large measurement uncertainty in TC of semiconductors, the logarithmic ($\log_{10}\kappa_L$) is usually employed in the absolute value of the performance assessment in the TC studies [28,39–41]. The average factor difference (AFD) [39], defined as $\mathrm{AFD} = 10^{x}$, where $x = \frac{1}{n}\sum_{n}\left|\log_{10}\kappa_{pred} - \log_{10}\kappa_{true}\right|$ (n is the number of data points employed), quantifies the average order of magnitude deviation of the model prediction from the ground truth. The DNN model prediction on the validation set is shown against the AGL data in Fig. 3a, and a high prediction accuracy of R^2 is 0.848 obtained with the corresponding AFD of 1.32 (equivalent mean absolute error (MAE) of $\log_{10}\kappa_L$ is 0.121). It should also be noted that R^2 values reported here are calculated after the TC values are converted to log scale, which we deem a fair way of characterizing the accuracy. The model parameters including the weights and biases of the DNNs trained on the AGL data are stored for further training.

TL for target task: For our target task, which is establishing ML model for the high-fidelity TC, we first directly train DNNs using 170 high-fidelity data with the same architecture as the proxy task (i.e., the same 91 shot models as for the AGL data). The five-fold cross-validation is adopted to evaluate the model performance on validation sets. The direct training gives an R^2 of 0.66 ± 0.06, and an AFD of 2.18 (Fig. 3b). We then apply TL in this target task, where we use the same DNN model architecture with the weights and biases from the DNN trained on the proxy AGL TC task as the starting parameters, and the parameters are directly trained. The DNN model is then re-trained using the small high-fidelity data set. As can be seen, TL can greatly improve R^2 and reduce the AFD. The R^2 can be improved by as much as 23% to 0.833, and the AFD can be reduced by as much as 30% to 1.52 (Fig. 3c). It should be noted that if the TC is transformed to a log scale before the TL approach, R^2 can be further improved, and the AFD can be further lowered (inset of Fig. 3d). Using log scales, however, has the assumption that the measurement uncertainty is proportional to the TC values, i.e., the high TC materials have larger uncertainty. For example, Raman spectroscopy is widely used for thermal conductivity measurement of 2D materials [42,43]. Due to the strong nonequilibrium of phonon modes, large uncertainty exists for the high TC graphene measurement [44]. By assuming the general proportionality of the same for all semiconductors, uncertainty of TC is then linearly proportional to its expected values, and thus applying log scales is a fair way to better account for the magnitude of the absolute uncertainty. It is also important to note that R^2 and AFD metrics are of great importance for the prediction of TC. These results suggest that TL leveraging AGL can indeed improve ML model accuracy, and the knowledge embedded in the AGL TC data set, though low-fidelity, can help optimize the ML model training on the small high-fidelity data set.

The performance of TL model can be influenced by the architecture and training/validation data partitions, and thus we show the plots of the R^2 and AFD metrics for all 91 shot cases in Fig. 4a, indicating that the increase in R^2 and

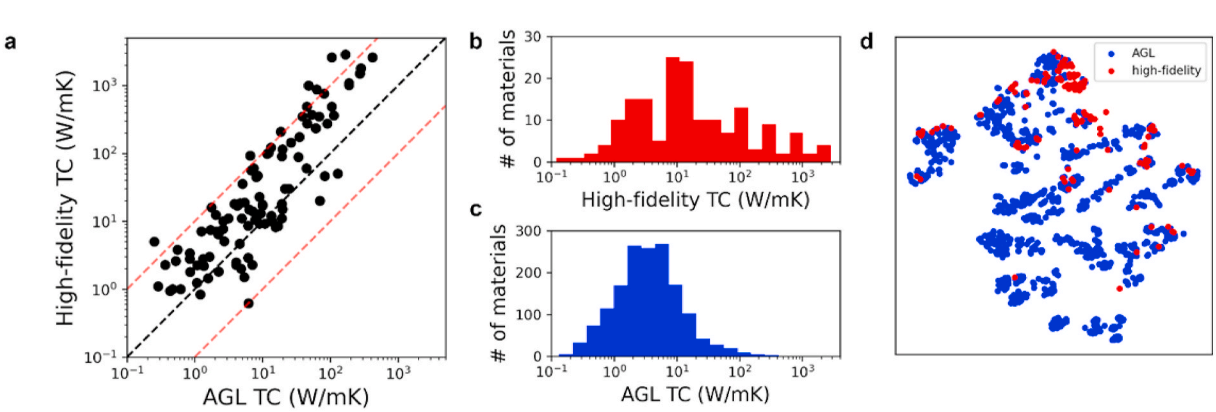

Fig. 2. Analysis of high-fidelity and low-fidelity data. (a) The comparison between the low-fidelity AGL TC data and the high-fidelity TC from experiments and first-principles calculations. AGL is sometimes underestimates the TC by an order of magnitude especially in the high TC region. (b) TC data distribution of the high-fidelity data set. (c) TC data distribution of the AGL data. (d) t-SNE visualization of the configurational descriptor space of the high-fidelity (red) and the AGL data (blue). The AGL data set covers much larger configurational space than the high-fidelity set.

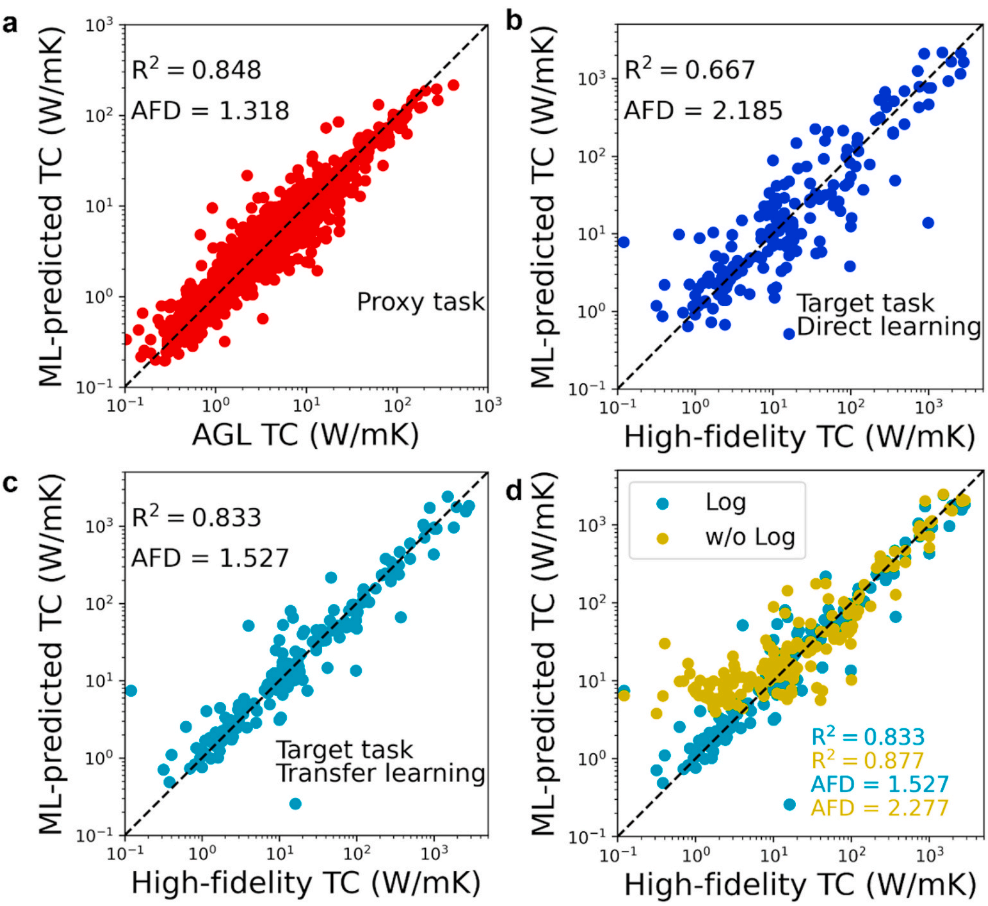

Fig. 3. ML model performance for different tasks. (a) Parity plot between ML-predicted TC and those in the AGL dataset. Note: All comparisons are based on the validation sets of five-fold validation. (b) (c) Parity plot between the ML-predicted TC and the high-fidelity TC in the validation sets where the ML model is (b) directly trained by the high-fidelity data, (c) from TL. (d) Parity plot between the ML-predicted TC and the high-fidelity TC using transfer learning with ('Log' in the legend) and without the log scale transformation ('w/o Log' in the legend).

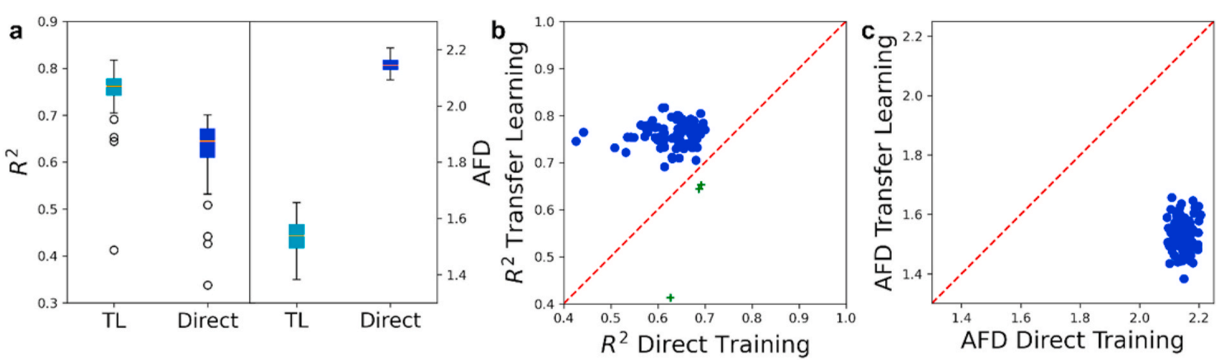

Fig. 4. Performance comparison between transfer learning (TL) model and direct learning model. (a) Box plot comparison of the R^2 and AFD of the two types of learning with the statistics from different cases. (b)(c) Scatter plot of the (b) R^2 and (c) AFD from two types of learning in all the cases.

decrease in AFD statistically. Fig. 4b shows the comparison of meaningful composition-property mapping learnt by the DNN of R^2 between direct learning and TL for all the 91 cases, which shows model. These observed effectiveness of the TL shows that TL can improve model accuracy in 97% of the cases (points above the diagonal in Fig. 4b). Fig. 4c shows the comparison of AFD, which indicates a successful meaningful composition-property TL can improve model accuracy in all the cases. In terms of R^2, the embedded AGL data and knowledge successfully majority of the TL model are above 0.7, while the direct learning ML model trained by the high-fidelity data directly hardly above 0.7. Similarly, in analyzing AFD over trained on the high-fidelity data, the prediction accuracy 2.0, while TL model can improve AFD to 1.4–1.6. Among the three validation settings, TL is still the most generalizable in all cases where TL does not outperform direct learning in R^2, two of them the trained model on the low prediction accuracy for validation have similar performance with only one case of using the biggest data and diverse configuration performance degradation from R^2 = 0.620 to 0.642 in different groups (Fig. 2c) of the AGL dataset. It is likely more universal model may capture different composition-property composition-property relationship which helps the TL model be more Considering the low-fidelity AGL TC and the high-fidelity TC generalizable. a good over the sample physics models, transferability. One of the design goals is the transferability to a potential application in mining candidate semiconductors that have high TC (e.g., >100 W/mK). A question arises if one should include all available data in the target task, including low TC values, in the TL model, or only need high TC values to achieve high accuracy for the high TC region. Another TL experiment is performed using the same DNN architecture for the source, where only TC > 100 W/mK data in the high-field dataset are used as the target of TL. The same AGL data is used as the proxy task, and the trained TL model (denoted as TL > 100) performance is evaluated using the model to predict TC in the validation set. As a comparison, we also test the previous TL and direct learning models to predict the TC of the same group of semiconductors and evaluate the prediction accuracy. The distribution of the performance difference is presented in Fig. 5. Although we expect the performance of TL using high-value field data can generally improve the models, TL based only on the high TC data are likely to yield less accurate models, it is still better than direct learning using all data. The observation suggests that low TC data also help for high TC prediction, as they may embed useful information about composition–property relationships. As a result, ultimately, it is desirable to leverage the full data as the target task to maximize the TL model accuracy.

Screening high TC semiconductors. We then use our trained TL model to identify high TC semiconductors by screening all the semiconductors in the Materials Project database [36], several materials are predicted to have high TC, and three examples are selected to validate using first-principles calculations (see Methods), considering their crystal structures. Two of these examples are GaB$_2$N (mp-100723) [45] and AlB$_2$N (mp-100857) [46], which has similar crystal structure with the space group number of 115. The crystal structures after optimization are shown in Fig. 6a. For GaB$_2$N, the TL-predicted TC is 540.25 ± 86.83 W/mK, while the first-principles calculated TC is 387.90 W/mK for the x-direction and 169 W/mK for the z-direction (Fig. 6b). For AlB$_2$N, the TL model predicted TC to be 533.73 ± 86.12 W/mK, while the first-principles TC is 510.30 W/mK along the x-direction and 164.06 W/mK in the z-direction. Another example tested is B$_2$AsP (mp-100828) [47], which has a tetragonal structure (Fig. 6a). The TL model predicted a TC of 611.77 ± 66.1 W/mK, while the first-principles calculated TC of 371.57 and 307.84 W/mK along the x- and z-directions, respectively. As we can see, while the absolute values predicted by the TL model are different from the first-principles results, they are able to predict TC in their right order of magnitude. For AlB$_2$N, the TL prediction is close to the ground truth in the x-direction. Characterizing the factor difference ($10^{|\log_{10}TC_{TL}-\log_{10}TC_{DFT}|}$), the errors of the three

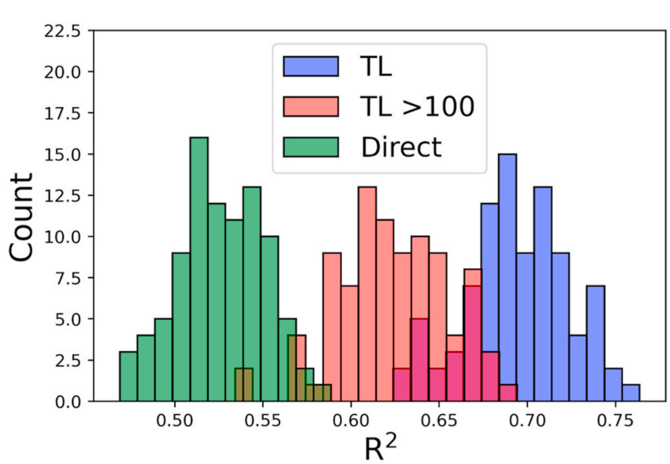

Fig. 5. TL strategy comparison for high TC prediction. The histograms of the performance coefficients of different models from direct learning, TL using large target task and TL using target task with TC > 100 W/mK.

compounds are 1.39 (GaB$_2$N), 1.05 (AlB$_2$N) and 1.64 (B$_2$AsP) for their x-direction TC, which is much lower than the AFD of the TL model (Fig. 6c). Of course, the TL models are trained using TC data without detailed information, understandably the predicted TC anisotropy only comes from the descriptors that consider anisotropic structural features. However, we believe the issues can be resolved with additional anisotropic data in the training sets, e.g., first-principles calculations indicate that BC$_2$N (mp-30148) is a good thermal conductor, with TC of around 910 W/mK [24]. Our TL model predicted a TC of 1075.86 ± 93.32 W/mK, agreeing well with the first-principles value. For the low TC semiconductors, experimental measurements for thermoelectric materials Y$_2$Te$_3$ and Ho$_2$Te$_3$ with room temperature TC values of 1.9 W/mK and 1.55 W/mK, respectively [40]. Our TL models predict TC values of 9.17 ± 8.39 W/mK and 6.69 ± 5.29 W/mK in their right order of magnitude.

3. Discussions

We note that a semi-empirical Debye–Callaway model is able to achieve AFD of ~1.5, which is better than our TL model, with the first-principles-based TC data [39] as the inputs. Debye temperature, speed of sound and Grüneisen parameters are calculated from density functional theory [48]. It should be noted that the fitting parameters in the model which usually need to be trained. In general, ad-hoc additional terms (e.g., structural parameters added) to the Debye–Callaway to improve model accuracy. The data-driven methods that no prior knowledge of the physics is needed, and they can take only material information (e.g., the chemistry, structure) as inputs, without the need for more complex descriptors from first-principles calculations. Of course, relying on data could lead to less accurate models, but our TL approach has been proven to solve such problem.

The ML model prediction performance sometimes significantly with the distribution of datasets used for training, validation and testing. It depends on how the performances are evaluated. These facts make fair comparison of models from different literature challenging. For example, Chen et al. [28] constructed a ML model using more sophisticated feature engineering and obtained a predictive $R^2 = 0.927$ on their test data. However, this metric was obtained from predicting a relatively small test data (25%) than ours, from uniformly random split of the absolute TC values. Our best performing model on some certain training/validation data splits could also achieve $R^2 = 0.97$. Zhu et al. [40] employed Crystal Graph Convolutional Neural Network (CGCNN) [49] to predict TC and achieved $R^2 = 0.85$ on a reserved test set, which is comparable to the R^2 in this work. However, we believe a model, which is the case for TC performance statistics by randomizing the training/validation split and DNN model as in our approach, should offer a more comprehensive sense of the TL performance. In addition, comparisons between TL and direct learning using the exact same set of models and training/validation data should reveal the true benefit of TL. The benefits should also be applicable to more sophisticated ML models like CGCNN [49].

The most important finding of this study is that we can use low-fidelity data as a proxy to perform TL in order to improve the ML model accuracy using small high-fidelity data. The low-fidelity data, although might contain larger errors, has valuable knowledge of composition–property relationship, which is transferable to the training with much smaller high-fidelity data. Due to the more diverse configurational space covered by the proxy data, the transferred composition–property relationship is helpful in improving the TL model when trained with the small high-fidelity data which cover much sparser configurational space. The implication of this study may be significant as it demonstrates the possibility

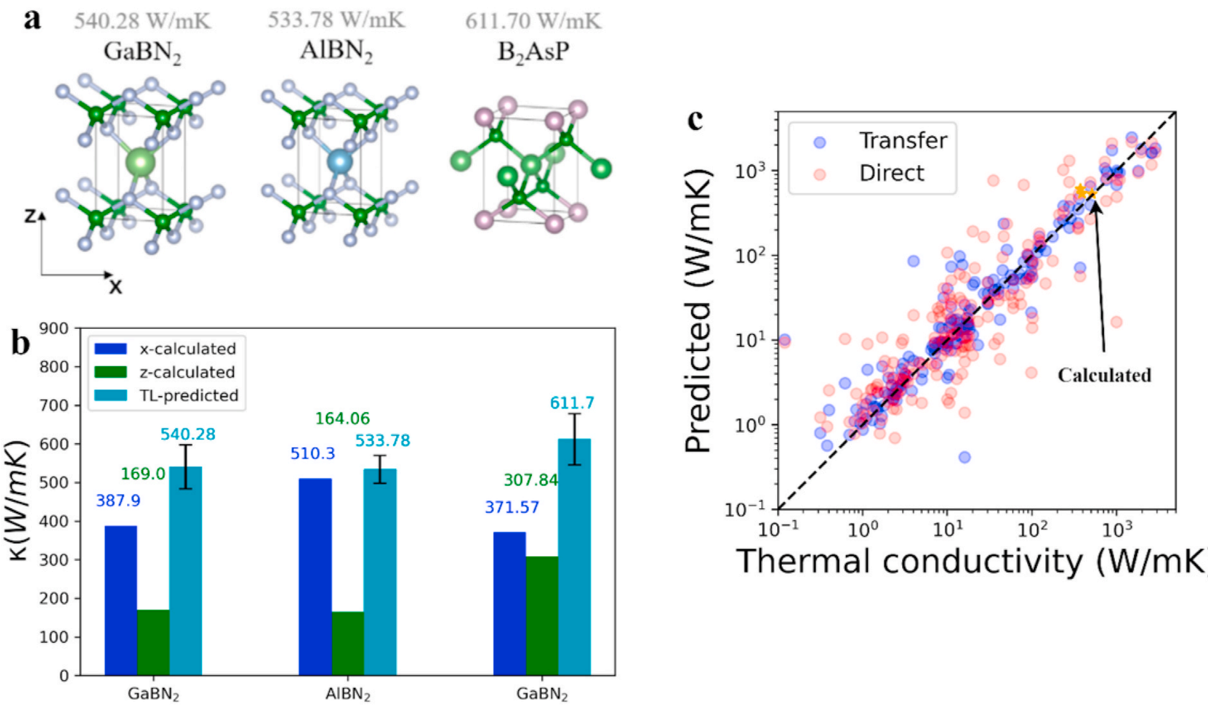

Fig. 6. Validation of TL-predicting TC of semiconductors. (a) Crystal structures of three example compounds selected for first-principles calculations. (b) Comparison plot of TL-predicted and first-principles TC values. (c) Calculated three compounds shown in the parity plot of TL-prediction on TC data.

using data from simplified empirical models can be generated in large volume as the proxy for the data-driven model trained with high-fidelity data that may be time- and resource-demanding to collect. Such a strategy could be beneficial to a wide range of applications in materials design.

4. Methods

Material representation: the present semiconductors are represented in a fixed length vector for ML tasks, the compositional descriptors implemented by the Matminer project are calculated using Magpie Data [56], where a total of 61 elemental property descriptors, such as atomic weight, atomic radius, etc., are included. Besides chemical descriptors, several structural descriptors including space group number, density, number of sites, and volume per unit cell, are also included, making the total dimension of the representation 165. As a standard ML practice, all features are scaled to the range of 0–1 during training. The high-fidelity thermal conductivity data is manually collected by searching the experimental and computational records of semiconductors at room temperature. And the low-fidelity thermal conductivity of semiconductors are collected from the Automatic-GIBBS-Library by restricting the band gap to be exist [21].

DNN for AGL TC: The DNN models are trained using PyTorch [38]. For the AGL TC (proxy property), there are three hidden layers in our DNN model connected with the ReLU activation function [51], and the number of neurons in each hidden layer is randomly selected. The first two layers in DNN contain random number of neurons (1st layer: 60–90 neurons, 2nd layer: 20–50 neurons), and 3rd layer contains fixed 10 neurons. A total of 91 such different models are pre-trained, and their performances are calculated using the ground truth property values by 15 independent five-fold cross-validation. The training epoch is set to be 500, with a 0.001 learning rate and the mini-batch size of 85. The model is then trained with Mean Squared Error (MSE) loss function with early stop. The evaluated prediction accuracy is the typical prediction of the validation data in the five-fold cross-validation processes are assembled and compared to ground truth. The R^2 and AFD as the metrics of the 91 models are averaged and the standard deviation is reported as the error bar.

Transfer learning (TL) in TC, we build DNN model with exactly the same architecture as the pre-trained proxy property model, parameters of the pre-trained models, except those of the output layer, are used as the initial parameters. These transferred models are then re-trained using the small data of the target task which builds the TL TC.

First-principles TC: For the first-principle lattice TC calculations, considering the three-phonon scattering processes, the iterative approach is employed to solve the phonon BTE to obtain the converged phonon distribution at 300 K using ShengBTE [52]. For the three-phonon calculation, a 6 × 6 × 6 q-grid is adopted. The density functional perturbation theory [53] scheme to calculate the harmonic force constants using Quantum Espresso [54]. The third order force constants are calculated using 4 × 4 × 4 supercell with the 8th nearest neighbor cutoff from finite-difference method. The isotope scattering based on the natural isotopic distribution is included using the Tamura equation [55]. The first Brillouin zone is discretized into a 12 × 12 × 12 grid, and the default Gaussian smearing is applied for the energy and crystal momentum conservation. These parameters are shown to yield converged TC values. It should be noted that for some materials with high temperature, higher-order phonon scattering like four-phonon scattering plays a role [18,56,57]. However, our first-principles calculations are to provide order of magnitude estimation of TC for comparison with TL model predictions. In addition, at the temperature of 300 K, we expect the effect of the four-phonon scattering to be relatively small.

Author credit

Zeyu Liu: Conceptualization, Methodology, Writing – original draft, Meng Jiang: Conceptualization, Supervision, Writing – review & editing. Tengfei Luo: Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tengfei Luo reports financial support was provided by National Science Foundation. Zeyu Liu reports financial support was provided by Fundamental Research Funds for the Central Universities.

Data availability

Data will be made available on request.

Acknowledgements

The computation is supported in part by the University of Notre Dame, Center for Research Computing, and NSF through XSEDE resources provided by TACC Stampede2 under grant number TG-CTS100078. Z.L. would like to thank the Fundamental Research Funds for the Central Universities (Grant No. 531118010723).

References

[1] Y. Cui, M. Li, Y. Hu, Emerging interface materials for electronics thermal management: experiments, modeling, and new opportunities, J. Mater. Chem. C 8 (2020) 10568.

[2] Q. Zheng, M. Hao, R. Miao, J. Schaadt, C. Dames, Advances in thermal conductivity for energy applications: a review, Prog. Energy 3 (2021) 012002.

[3] Z. Tian, S. Lee, G. Chen, Comprehensive review of heat transfer in thermoelectric materials and devices, Annu. Rev. Heat Transf. 17 (2014).

[4] E.S. Toberer, L.L. Baranowski, C. Dames, Advances in thermal conductivity, Annu. Rev. Mater. Res. 42 (2012) 179.

[5] R. Daro, Thermal interface materials: technology review, update, remaining challenges, Mater. Res. 58 (2013) 15.

[6] C. Chiritescu, D.G. Cahill, N. Nguyen, D. Johnson, A. Bodapati, P. Keblinski, P. Zschack, Ultralow thermal conductivity in disordered, layered WSe2 crystals, Science 315 (2007) 351.

[7] S. Kim, et al., Extremely anisotropic van der Waals thermal conductors, Nature 597 (2021) 660.

[8] S. Mukhopadhyay, D.S. Parker, B.C. Sales, A.A. Puretzky, M.A. McGuire, L. Lindsay, Two-channel model for ultralow thermal conductivity of crystalline Tl3VSe4, Science 360 (2018) 1455.

[9] S.-F. Wang, Z.-G. Zhang, B.-T. Wang, J.-R. Zhang, F.-W. Wang, Intrinsic ultralow lattice thermal conductivity in the full-Heusler compound Ba2AgSb, Phys. Rev. Appl. 17 (2022) 034023.

[10] J.S. Kang, M. Li, H. Wu, H. Nguyen, Y. Hu, Experimental observation of high thermal conductivity in boron arsenide, Science 361 (2018) 575.

[11] F. Tian, et al., Unusual high thermal conductivity in boron arsenide bulk crystals, Science 361 (2018) 582.

[12] S. Li, Q. Zheng, Y. Lv, X. Liu, X. Wang, P.Y. Huang, D.G. Cahill, B. Lv, High thermal conductivity in cubic boron arsenide crystals, Science 361 (2018) 579.

[13] K. Chen, et al., Ultrahigh thermal conductivity in isotope-enriched cubic boron nitride, Science 367 (2020) 555.

[14] L. Lindsay, D.A. Broido, T.L. Reinecke, First-principles determination of ultrahigh thermal conductivity of boron arsenide: a competitor for diamond? Phys. Rev. Lett. 111 (2013) 025901.

[15] A.J.H. McGaughey, A. Jain, H.-Y. Kim, B. Fu, Phonon properties and thermal conductivity from first principles, lattice dynamics, and the Boltzmann transport equation, J. Appl. Phys. 125 (2019) 011101.

[16] L. Lindsay, First principles Peierls-Boltzmann phonon thermal transport: a topical review, Nanoscale Microscale Thermophys. Eng. 20 (2016) 67.

[17] K. Esfarjani, G. Chen, H.T. Stokes, Heat transport in silicon from first-principles calculations, Phys. Rev. B 84 (2011) 085204.

[18] T. Feng, L. Lindsay, X. Ruan, Four-phonon scattering significantly reduces intrinsic thermal conductivity of solids, Phys. Rev. B 96 (2017) 161201.

[19] J.J. Plata, P. Nath, D. Usanmaz, J. Carrete, C. Toher, M. de Jong, M. Asta, M. Fornari, M.B. Nardelli, S. Curtarolo, An efficient and accurate framework for calculating lattice thermal conductivity of solids: AFLOW—AAPL automatic anharmonic phonon library, Npj Comput. Mater. 3 (2017) 45.

[20] G.A. Slack, Nonmetallic crystals with high thermal conductivity, J. Phys. Chem. Solid. 34 (1973) 321.

[21] C. Toher, J.J. Plata, O. Levy, M. de Jong, M. Asta, M.B. Nardelli, S. Curtarolo, High-throughput computational screening of thermal conductivity, Debye temperature, and Grüneisen parameter using a quasiharmonic Debye model, Phys. Rev. B Condens. Matter 90 (2014).

[22] C. Sutton, L.M. Ghiringhelli, T. Yamamoto, Y. Lysogorskiy, L. Blumenthal, T. Hammerschmidt, J.R. Golebiowski, X. Liu, A. Ziletti, M. Scheffler, Crowd-sourcing materials-science challenges with the NOMAD 2018 Kaggle competition, Npj Comput. Mater. 5 (2019).

[23] G. Pilania, J.E. Gubernatis, T. Lookman, Multi-fidelity machine learning models for accurate bandgap predictions of solids, Comput. Mater. Sci. 129 (2017) 156.

[24] S. Ju, R. Yoshida, C. Liu, S. Wu, K. Hongo, T. Tadano, J. Shiomi, Exploring diamondlike lattice thermal conductivity crystals via feature-based transfer learning, Phys. Rev. Mater. 5 (2021) 053801.

[25] T. Wang, C. Zhang, H. Snoussi, G. Zhang, Machine learning approaches for thermoelectric materials research, Adv. Funct. Mater. 30 (2020) 1906041.

[26] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization, Phys. Rev. Lett. 115 (2015) 205901.

[27] J. Carrete, W. Li, N. Mingo, S. Wang, S. Curtarolo, Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling, Phys. Rev. X 4 (2014) 011019.

[28] L. Chen, H. Tran, R. Batra, C. Kim, R. Ramprasad, Machine learning models for the lattice thermal conductivity prediction of inorganic materials, Comput. Mater. Sci. 170 (2019) 109155.

[29] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345.

[30] S. Kong, D. Guevarra, C.P. Gomes, J.M. Gregoire, Materials representation and transfer learning for multi-property prediction, Appl. Phys. Rev. 8 (2021) 021409.

[31] H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, R. Yoshida, Predicting materials properties with little data using shotgun transfer learning, ACS Cent. Sci. 5 (2019) 1717.

[32] Z. Liu, Jiang Meng, and Tengfei Luo, Leverage electron properties to predict phonon properties via transfer learning for semiconductors, Sci. Adv. 6 (2020) eabd1356 (n.d.).

[33] R. Ma, Y.J. Colón, T. Luo, Transfer learning study of gas adsorption in metal–organic frameworks, ACS Appl. Mater. Interfaces 12 (2020) 34041.

[34] D. Jha, K. Choudhary, F. Tavazza, W. Liao, A. Choudhary, C. Campbell, A. Agrawal, Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning, Nat. Commun. 10 (2019) 5316.

[35] J.P. Perdew, Density functional theory and the band gap problem, Int. J. Quant. Chem. 28 (1985) 497.

[36] A. Jain, et al., Commentary: the materials project: a materials genome approach to accelerating materials innovation, Apl. Mater. 1 (2013) 011002.

[37] F. Pedregosa, et al., Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825.

[38] A. Paszke, et al., PyTorch: an imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035.

[39] S.A. Miller, et al., Capturing anharmonicity in a lattice thermal conductivity model for high-throughput predictions, Chem. Mater. 29 (2017) 2494.

[40] T. Zhu, R. He, S. Gong, T. Xie, P. Gorai, K. Nielsch, J.C. Grossman, Charting lattice thermal conductivity for inorganic crystals and discovering rare earth chalcogenides for thermoelectrics, Energy Environ. Sci. 14 (2021) 3559.

[41] R. Jaafreh, Y.S. Kang, K. Hamad, Lattice thermal conductivity: an accelerated discovery guided by machine learning, ACS Appl. Mater. Interfaces 13 (2021) 57204.

[42] A.A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, C.N. Lau, Superior thermal conductivity of single-layer graphene, Nano Lett. 8 (2008) 902.

[43] R. Yan, J.R. Simpson, S. Bertolazzi, J. Brivio, M. Watson, X. Wu, A. Kis, T. Luo, A.R. Hight Walker, H.G. Xing, Thermal conductivity of monolayer molybdenum disulfide obtained from temperature-dependent Raman spectroscopy, ACS Nano 8 (2014) 986.

[44] A.K. Vallabhaneni, D. Singh, H. Bao, J. Murthy, X. Ruan, Reliability of Raman measurements of thermal conductivity of single-layer graphene due to selective electron-phonon coupling: a first-principles study, Phys. Rev. B 93 (2016) 125432.

[45] K. Persson, Materials Data on GaBN2 (SG:115) by Materials Project.

[46] K. Persson, Materials Data on AlBN2 (SG:115) by Materials Project.

[47] M. Merabet, D. Rached, R. Khenata, S. Benalia, B. Abidri, N. Bettahar, S. Bin Omran, Electronic structure of (BP)n/(BAs)n (001) superlattices, Physica B Condens. Matter 406 (2011) 3247.

[48] J. Yan, P. Gorai, B. Ortiz, S. Miller, S.A. Barnett, T. Mason, V. Stevanović, E.S. Toberer, Material descriptors for predicting thermoelectric performance, Energy Environ. Sci. 8 (2015) 983.

[49] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. 120 (2018) 145301.

[50] L. Ward, et al., Matminer: an open source toolkit for materials data mining, Comput. Mater. Sci. 152 (2018) 60.

[51] M.D. Zeiler, et al., On rectified linear units for speech processing, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 3517–3521.

[52] W. Li, J. Carrete, N.A. Katcho, N. Mingo, ShengBTE: a solver of the Boltzmann transport equation for phonons, Comput. Phys. Commun. 185 (2014) 1747.

[53] S. Baroni, S. de Gironcoli, A. Dal Corso, P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, Rev. Mod. Phys. 73 (2001) 515.

[54] P. Giannozzi, et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys. Condens. Matter 21 (2009).

[55] S. Tamura, Isotope scattering of large-wave-vector phonons in GaAs and InSb: deformation-dipole and overlap-shell models, Phys. Rev. B 30 (1984) 849.
[56] T. Feng, X. Ruan, Quantum mechanical prediction of four-phonon scattering rates and reduced thermal conductivity of solids, Phys. Rev. B 93 (2016) 045202.
[57] Z. Han, X. Yang, W. Li, T. Feng, X. Ruan, FourPhonon: an extension module to ShengBTE for computing four-phonon scattering rates and thermal conductivity, Comput. Phys. Commun. 270 (2022) 108179.