Exploring diamondlike lattice thermal conductivity crystals via feature-based transfer learning

Shenghong Ju
Department of Mechanical Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan and National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

Ryo Yoshida
Research Organization of Information and Systems, The Institute of Statistical Mathematics (ISM), 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan and National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

Chang Liu and Stephen Wu
Research Organization of Information and Systems, The Institute of Statistical Mathematics (ISM), 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

Kenta Hongo
Japan Advanced Institute of Science and Technology (JAIST), 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan; National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan; and PRESTO, Japan Science and Technology Agency (JST), Kawaguchi, Saitama 332-0012, Japan

Terumasa Tadano
National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

Junichiro Shiomi [*]
Department of Mechanical Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan; National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan; Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), 4-1-8, Kawaguchi, Saitama 332-0012, Japan; and RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

(Received 15 April 2018; revised 7 January 2019; accepted 16 April 2021; published 10 May 2021)

Ultrahigh lattice thermal conductivity materials hold great importance since they play a critical role in the thermal management of electronic and optical devices. Models using machine learning can search for materials with outstanding higher-order properties like thermal conductivity. However, the lack of sufficient data to train a model is a serious hurdle. Herein we show that big data can complement small data for accurate predictions when lower-order feature properties available in big data are selected properly and applied to transfer learning. The connection between the crystal information and thermal conductivity is directly built with a neural network by transferring descriptors acquired through a pretrained model for the feature property. Successful transfer learning shows the ability of extrapolative prediction and reveals descriptors for lattice anharmonicity. The resulting model is employed to screen over 60 000 compounds to identify novel crystals that can serve as alternatives to diamond. Even though most materials in the top list are superhard materials, we reveal that superhard property does not necessarily lead to high lattice thermal conductivity. Large hardness means high elastic constants and group velocity of phonons in the linear dispersion regime, but the lattice thermal conductivity is determined also by other important factors such as the phonon relaxation time. What is more, the average or maximum dipole polarizability and the van der Waals radius are revealed to be the leading descriptors among those that can also be qualitatively related to anharmonicity.

DOI: 10.1103/PhysRevMaterials.5.053801

I. INTRODUCTION

The power densities of microelectronic devices and their components continually increase due to advances in the fabrication and integration of advanced materials and structures. Hence, the large thermal density must be quickly removed to guarantee reliable performance. Material innovations in heat spreaders and sinks and thermal interface materials are at the core of the thermal-management challenge. A key element to such innovations is materials with a high lattice thermal conductivity (κ_L) either as bulk crystals or fillers for composites.

[*]To whom correspondence should be addressed: shiomi@photon.t.u-tokyo.ac.jp

Although metals are generally suitable thermal conductors, insulators have the highest thermal conductivities. In many thermal management applications involving heat spreaders and sinks, electrical insulation is necessary to avoid electric current leakage. Diamond, which has a thermal conductivity of about 2000 $\mathrm{Wm^{-1}\,K^{-1}}$ at room temperature, is a representative bulk material [1]. It is widely used as heat spreaders or sinks for laser diodes and power electronics in the form of bulk or composites to prevent overheating. One drawback is that it sustains thermal damage via oxidation or graphitization at high temperature [2,3], significantly altering the thermal properties of the heat spreader or sink.

Cubic and hexagonal boron nitrides have been investigated as alternative materials [4]. Considering the surface affinities with various other materials for composite syntheses and integration, other alternatives should be useful. Although physics insights suggest that some materials exhibit a fairly high κ_L such as SiC, BeO, BP, AlN, BeS, GaN, Si, AlP, and GaP, materials with κ_L approaching or exceeding 1000 $\mathrm{Wm^{-1}\,K^{-1}}$ are rare.

Single-crystal compounds are obvious candidates as alternative high-κ_L materials to diamond [5]. However, only a few materials have quantified thermal conductivity values due to the difficulty of synthesizing single crystals that can be measured in a standardized fashion. Moreover, a material search is extremely cumbersome. Herein we propose utilizing computational techniques to efficiently search for high-κ_L materials. In recent decades, the development of lattice dynamics methods using interatomic force constants obtained from density functional theories has enabled first-principles calculations of the κ_L. Simultaneously, databases containing tens of thousands of crystal compounds have been constructed. Examples include MATERIALS PROJECT [6], AFLOW [7], ICSD [8], and ATOMWORK [9]. However, performing first-principles calculations for all the crystals in the databases is extremely time consuming and unrealistic.

Another option is high throughput screening [10] based on machine learning. High throughput screening can speed up discovery of new materials. It has been applied in many fields such as catalysis, battery technologies, thermoelectric materials, chemical probes, polymers, and magnetic materials. Motivated by realizing high-performance thermoelectric materials, efforts to apply κ_L to crystals have centered on screening ultralow κ_L crystals [11–13]. Carrete *et al.* [13] screened 79 000 half-Heusler compounds and found that materials with large atomic radii elements have a lower κ_L. Seko *et al.* [12] screened 54 779 crystals based on the Gaussian process regression and reported 221 materials with a low κ_L. Roekeghem *et al.* [11] extended the screening of mechanically stable compounds at high temperatures using finite-temperature phonon calculations.

One challenge when screening crystals with high and low thermal conductivities is the large gap between the "big data" required for credible machine learning and the "small data" currently available. Although bridging this gap is a general problem in materials informatics [14–16], it is especially intense when searching for materials with a preferred thermal conductivity because it involves both the harmonic phonon properties, which are fairly easy to calculate, and anharmonic properties, which are much more expensive to evaluate. As shown in Fig. 1, "big data" are available for the harmonic phonon property of the three-phonon scattering phase space but not for the thermal conductivity (Fig. 1). Only "small data" are available for the thermal conductivity due to the heavy calculation required for the anharmonic phonon property. To overcome the difficulty of limited data availability, we employ a transfer learning strategy which is an increasingly popular technique of machine learning [17]. Because openly accessible big data are less available in material research, the ability of transfer learning to learn on small data has attracted much attention [18–20]. Yamada *et al.* [21] demonstrated the potential power of transfer learning in various applications in materials science, including organic and inorganic chemistry. In this study, we show that a prediction model with fairly high accuracy can be derived from the limited data set by effectively reusing features learned on the harmonic phonon property as features in the model of the anharmonic phonon property.

In this work, we develop a feature-based transfer learning method to overcome the gap. This method begins with a broad search over the entire structure database of crystal compounds for the feature harmonic property, which should be correlated with thermal conductivity. Subsequently, a focused search of the selected candidates is conducted for a high thermal conductivity. Here, we choose the scattering phase space (P_3) of the three-phonon scattering process as the feature property because this can be quickly extracted from the harmonic calculation.

II. METHODS

A. Anharmonic lattice dynamics

The conventional method of calculating κ_L is formulated on the basis of anharmonic lattice dynamics calculations using the interatomic force constants obtained by first principles, which are well described elsewhere [22]. The κ_L was calculated by solving the Boltzmann transport equation (see Supplemental Material, Note A [23])

$$-\boldsymbol{v}_{\boldsymbol{q}s}\dot{\nabla}T\left(\frac{\partial n_{\boldsymbol{q}s}}{\partial T}\right)+\left.\frac{\partial n_{\boldsymbol{q}s}}{\partial t}\right|_{\text{scattering}}=0, \tag{1}$$

where n is the phonon distribution function, $\boldsymbol{q}s$ is the phonon mode, and $\boldsymbol{v}$ is the group velocity. Although the single-mode relaxation time approximation (see Supplemental Material, Note B [23]) is often used to calculate κ_L, it significantly underestimates κ_L for high-thermal-conductivity crystals. Thus, Eq. (1) must be solved iteratively or directly (see Supplemental Material, Fig. S2 and Note C [23]).

As discussed above, a key feature in the current process is P_3, which quantifies the phonon scattering channels. The total P_3 is calculated as

$$P_3=\frac{1}{N_q}\sum_{\boldsymbol{q}s}\frac{1}{3m^3}\left(2P_3^{(+)}(\boldsymbol{q}s)+P_3^{(-)}(\boldsymbol{q}s)\right), \tag{2}$$

where m is the number of phonon branches and

$$P_3^{(\pm)}(\boldsymbol{q}s)=\frac{1}{N_q}\sum_{\boldsymbol{q}'s',\boldsymbol{q}''s''}\delta(\omega_{\boldsymbol{q}s}\pm\omega_{\boldsymbol{q}'s'}-\omega_{\boldsymbol{q}''s''})\delta_{\boldsymbol{q}\pm\boldsymbol{q}',\boldsymbol{q}''+\boldsymbol{G}}. \tag{3}$$

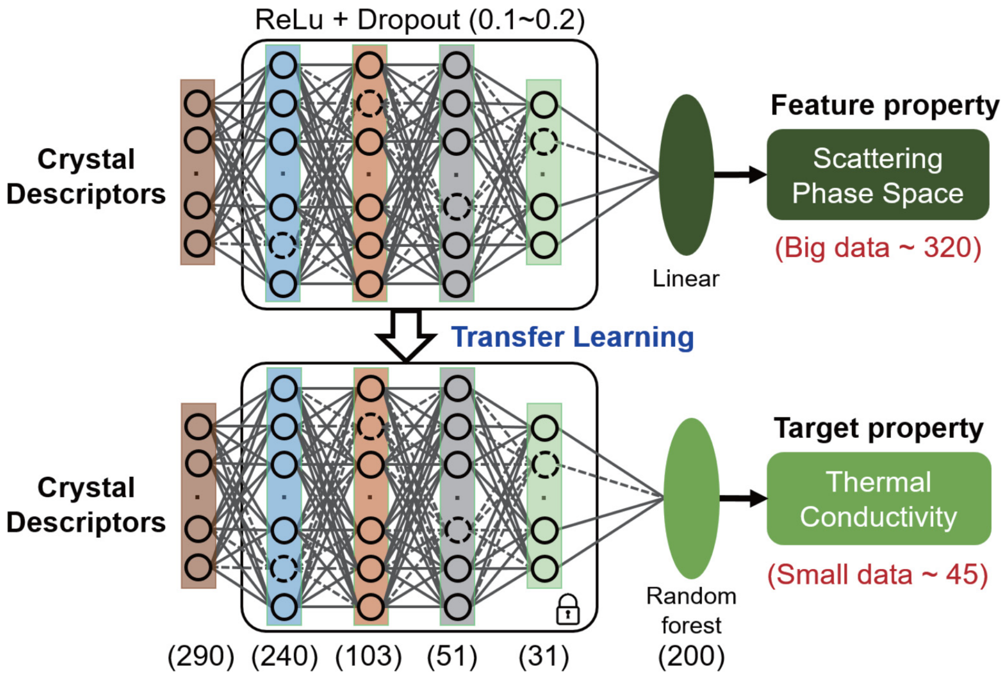

FIG. 1. Schematics of feature-based transfer learning. The transfer learning bridges "big data" (harmonic three-phonon scattering phase space of 320 crystals) and "small data" (thermal conductivity of 45 crystals) to search for ultrahigh lattice thermal conductivity crystals. All neurons (circles) are activated by ReLU (rectified linear unit). Dropout (dashed circles and lines) range (0.1 or 0.2) in each hidden layer is randomly chosen. Numbers at the bottom indicate the number of neurons or trees used in each layer of the neural network and random forest model.

Equation (3) indicates that P_3 can be calculated solely from harmonic interatomic force constants.

B. P_3 and κ_L data collections

The candidates were approximately 60 000 inorganic crystals in the MATERIALS PROJECT [6] database. Because this study focused on κ_L, materials with a band gap smaller than 0.1 eV, molecular crystals such as O_2, H_2, H_2O, H_2O_2, and crystals with hydrogen atoms were excluded. We collected the atom displacement from the phonon database [24] and calculated force data of 320 crystals by first principles (see Supplemental Material, Table S1 [23]). The ALAMODE package [25] was then used to fit the harmonic interatomic force constants and calculate the P_3 values. Aside from the P_3 data, we also collected thermal conductivity data for 45 materials (see Supplemental Material, Table S2 [23]).

C. Transfer learning

For the given target property (κ_L), which has limited training data, models on the proxy feature property (P_3) are pretrained using the sufficient data to capture the features relevant to the commonality between κ_L and P_3. Repurposing the pretrained feature extractor on the target task can realize an outstanding prediction ability even with the exceedingly small amount of data. This study focuses on a specific type of transfer learning using artificial neural networks (Fig. 1).

Transfer learning was performed via XENONPY, a self-developed open-source Python package [26]. We used XENONPY to calculate 290 compositional features for a given chemical composition using information about the 58 per-element features, such as the atomic weight, electronegativity, van der Walls radius, and so on. We pretrained a fully connected pyramid neural network using the 320 instances of P_3 and the 290-dimensional descriptor vectors. All neurons were activated by ReLU (rectified linear unit), and a linear model was placed on the output layer, which defines the transformation from the 10 neurons in the last hidden layer to the P_3. We produced 1000 pretrained models on the P_3 with randomly generated network structure; the number of hidden layers, which ranged between 4 and 6; the number of neurons in each layer; and the dropout probability, which was either 0.1 or 0.2. Subsequently, the best model on the P_3 was selected based on the 10-fold cross validation looped within the 320 instances. Except for the output layer, the subnetwork of the selected model was used as both a feature extractor and an input descriptor in the prediction model of the κ_L. Finally, the random forest (the number of trees = 200) model was selected and trained using the 45 instances (see Supplemental Material, Table S2 [23]) of the κ_L and the 10-dimensional descriptors acquired through the pretraining process.

III. RESULTS AND DISCUSSION

A. Performance of transfer learning

The prediction model connecting the input materials and P_3 was initially trained based on the 320 collected instances of P_3 and 290 descriptors. The subnetwork of the pretrained model was transferred to train the model connecting the crystal structures and κ_L by replacing the linear output layer with the random forest model using the 45 κ_L data. The training and validation results of the pretrained and transferred models are shown in Figs. 2(a) and 2(b).

Using the models for the κ_L and P_3 in the high-throughput screening with ~60 000 crystals in MATERIALS PROJECT[6], we identified the top-14 crystals with the smallest P_3 (see Supplemental Material, Table S3 [23]) from the top-100 prediction list whose κ_L were validated using the first-principles based anharmonic lattice dynamics calculations (Fig. 3 and Table I). The calculation details can be seen in the Supplemental Material, Table S7 and Note D [23]. The transferred model successfully predicts the 14 crystals even though their κ_L lie in the ultrahigh region of 1000–3000 $\mathrm{Wm^{-1}\,K^{-1}}$ [Fig. 2(c)]. It should be noted that the κ_L of the 45 training crystals reside in the region smaller than 370 $\mathrm{Wm^{-1}\,K^{-1}}$, which is much lower than the prediction [Fig. 2(d)]. This indicates that the transferred model exhibits "extrapolative predictive power." In general, ordinary machine learning is "interpolative," and its prediction ability is applicable only in a neighboring region of the given training instances.

Indeed, a neural network directly trained using the 45 samples performs rather poorly for the 14 crystals as the predicted κ_L never exceeded 600 $\mathrm{Wm^{-1}K^{-1}}$ [Fig. 2(c) and see Supplemental Material, Table S4 [23]]. The pretraining process using the 320 instances on P_3 would contribute to the acquisition of the extrapolative ability. The pretrained neural network on P_3 is able to represent material structures that are applicable to a broader input space than the one spanned by much fewer instances of κ_L. This is because the 320 source P_3 data contain instances that account for the structure-property relationships relevant to the ultrahigh thermal conductivity.

B. Top high-thermal-conductivity crystals

We identified the top-14 materials by feature-based transfer learning. They are comprised of boron arsenides

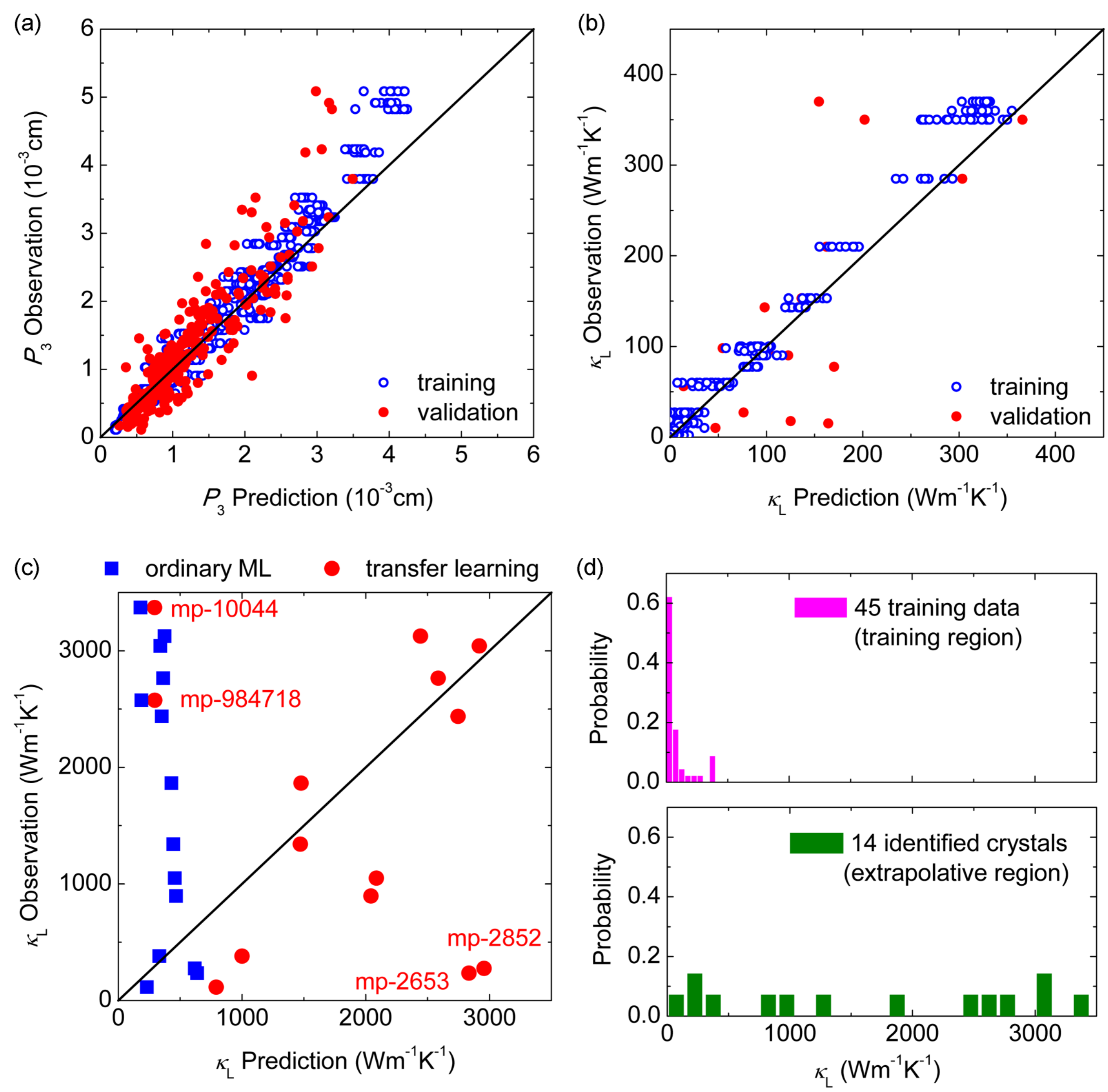

FIG. 2. Performance of transfer learning. Training and validation for (a) the pretrained P_3 model (mean absolute error = 0.000 237 cm) and (b) transferred κ_L model (mean absolute error = 30.8528 $\mathrm{Wm^{-1}\,K^{-1}}$). Hollow and solid dots denote the results of training and testing in the cross validation for the best prediction model. (c) Comparison between ordinary machine learning and transfer learning. (d) κ_L distribution of 45 training and 14 identified crystals.

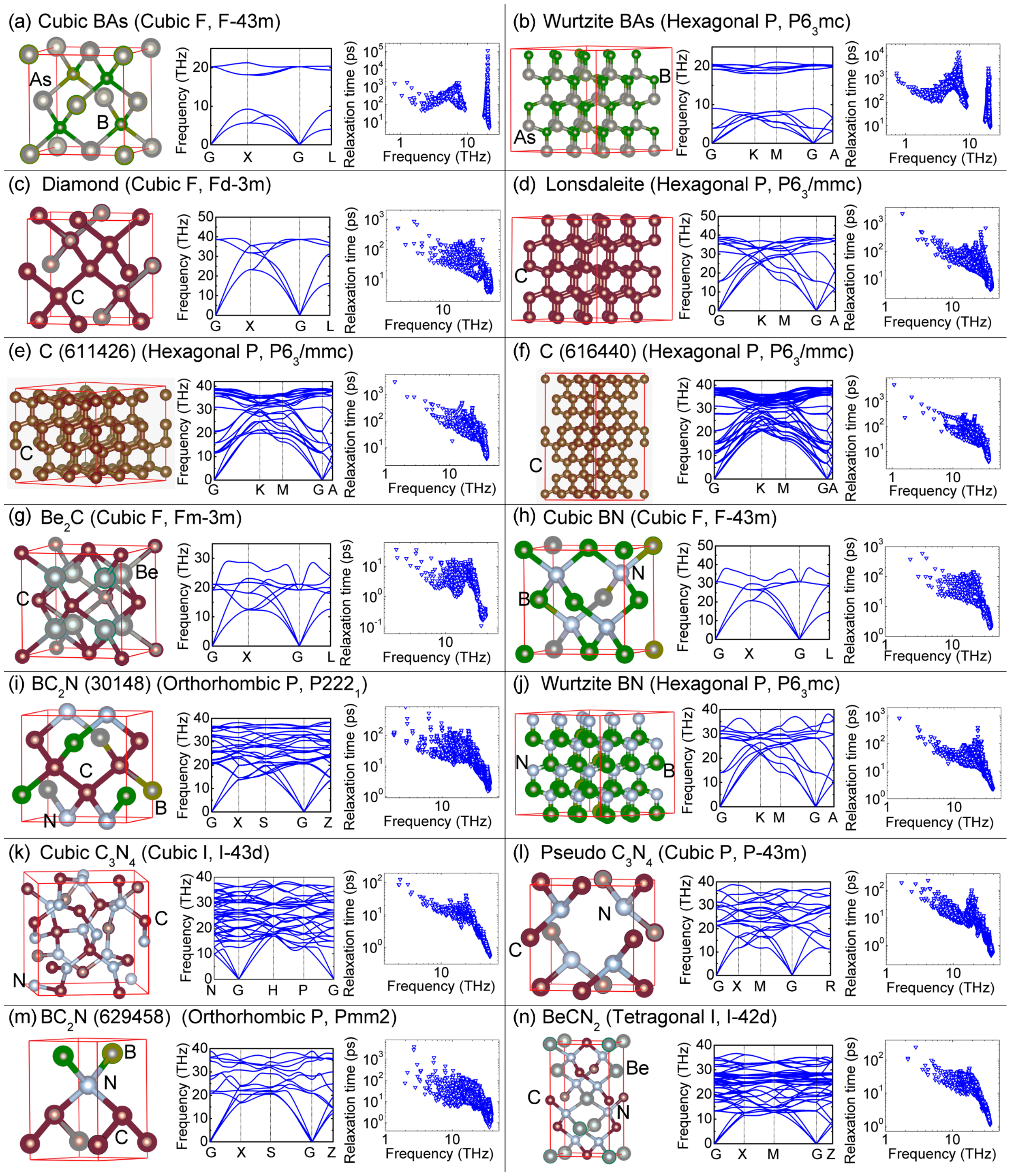

FIG. 3. Top-14 crystals with the lowest calculated P_3. Crystal structures, phonon dispersions, and phonon relaxation times of the top-14 materials with the lowest calculated P_3. Parentheses indicate the crystal structure.

(BAs), carbon (C), beryllium carbide (Be_2C), boron nitride (BN), heterodiamond (BC_2N), carbon nitride (C_3N_4), and ternary $BeCN_2$. They all have thermal conductivities above 100 $Wm^{-1}K^{-1}$. In fact, 10 exceed 500 $Wm^{-1}K^{-1}$. The top two crystals with the lowest calculated P_3 are cubic and wurtzite BAs, which have thermal conductivity over 1000 $Wm^{-1}K^{-1}$.

Recently, many studies have investigated cubic BAs due to their high predicted thermal conductivity (3170 $Wm^{-1}K^{-1}$) [27], which is comparable with diamond. However, these initial experiments measured the thermal conductivity of BAs crystals around 186–350 $Wm^{-1}K^{-1}$ [28,32,33]. This difference is attributed mainly to the difficulty in the fabrication of single crystals of boron-related materials as well as the complicated synthesis due to the high volatility and toxicity of arsenide atoms. Although the four-phonon scattering process is important for high-thermal-conductivity materials at high temperatures [34], the theoretical thermal conductivity

TABLE I. Top-14 materials with the lowest P_3 values and their thermal conductivities calculated by the iterative Boltzmann transport equation solution. *xx*, *yy*, and *zz* indicate the lattice directions.

			Thermal conductivity ($Wm^{-1}K^{-1}$)				
		P_3		This work		Calc. Ref.	Expt. Ref.
Name	Structure	(10^{-4} cm)	*xx*	*yy*	*zz*	*xx*/*yy* (*zz*)	*xx*/*yy* (*zz*)
Cubic BAs	*F*-43*m*	0.6397	3411	3411	3411	3170 [27]	351 [28]
Wurtzite BAs	*P*63*mc*	0.9064	2947	2947	1881	2380 (1210) [29]	
Diamond	*Fd*-3*m*	1.0005	3048	3048	3048	3450 [27]	3000 [1]
Lonsdaleite	*P*63/*mmc*	1.0335	2533	2533	2122	1500 (1270) [30]	
C (611426)	*P*63/*mmc*	1.2437	2842	2842	2675		
C (616440)	*P*63/*mmc*	1.2569	2583	2583	4214		
Be_2C	*Fm*-3*m*	1.2596	117	117	117		
Cubic BN	*F*-43*m*	1.3300	1876	1876	1876	1800 [29]	768 [31]
BC_2N (30148)	*P*2221	1.3670	895	910	804		
Wurtzite BN	*P*63*mc*	1.4394	1359	1359	1305	1230 (1040) [29]	
Cubic C_3N_4	*I*-43*d*	1.4490	234	234	234		
Pseudo C_3N_4	*P*-43*m*	1.4529	275	275	275		
BC_2N (629458)	*Pmm*2	1.5127	1392	972	784		
$BeCN_2$	*I*-42*d*	1.5472	351	351	440		

of BAs remains as high as 2000 $Wm^{-1}K^{-1}$. In fact, recent experiments realized cubic BAs with a thermal conductivity as high as 1000 $Wm^{-1}K^{-1}$ [35–37].

Cubic BAs should be a reasonable demonstration of the effectiveness of the current screening. In this paper, we focus on the thermal conductivity at room temperature. Because the three-phonon scattering rate is much higher than the four-phonon scattering rate, employing three-phonon P_3 is reasonable to find high-thermal-conductivity materials.

Another interesting feature of the top list is that it contains allotropes of known high-thermal-conductivity cubic materials (diamond, BN, and BAs). These include lonsdaleite, hexagonal diamonds, wurtzite BN, and wurtzite BAs, which have not been studied previously in terms of thermal conductivity. These materials are potential alternatives to their cubic counterparts. Although both cubic BN and wurtzite BN can be formed by compressing hexagonal BN, wurtzite BN is formed at much lower temperatures (around 2000 K) than cubic ones (3000–4000 K) [38]. Recently, a single-phase wurtzite BN bulk crystal was synthesized directly from a hexagonal BN bulk crystal under 10 GPa and 850 °C [39]. If the wurtzite structure has similar merits in other species, then BAs may be the first to benefit.

The list of top-100 materials with the smallest P_3 also contains some typical high-thermal-conductivity materials like SiC (mp-ID: 8062), GaN (mp-ID: 830), and AlN (mp-ID: 1700) (see Supplemental Material, Table S3 [23]). Additionally, the top-100 list includes different crystal structures of GaN (mp-ID: 804) and AlN (mp-ID: 1330), which may display high thermal conductivities. Layered structure materials such as hexagonal BN and graphite have high thermal conductivities in the in-plane direction, but their out-of-plane thermal conductivity is very low due to the weak atomistic interaction. Hexagonal BN (mp-ID: 984) is in the top-100 prediction list as well as other BNs with layered structures. Examples include mp-ID: 7991, 685145, 13150, 604884, 629015, and 569655. Moreover, the list has some graphite structures (mp-ID: 568806, 632329, 990448, 568286 and 569304). In this work, we evaluate the thermal conductivity with the scattering phase space, which is a scalar parameter that tends to suggest crystals with a high thermal conductivity in all three-lattice directions. The experimental conditions are the main reason why two-dimensional (2D) materials such as hexagonal BN and graphite do not appear in our top-14 list.

C. Hardness versus thermal conductivity

An interesting feature of the screening results is that most of the top-14 list are superhard materials, including diamond, carbon nitride, born nitride, and heterodiamond. The Vickers hardness of diamond is around 115 GPa [40], which is the highest among reported superhard crystals, including cubic BN (62 GPa) [41], cubic BC_2N (76 GPa) [40], C_3N_4 (37–90 GPa) [42], and $BeCN_2$ (37 GPa) [43]. Be_2C has a Knoop hardness of 2410 kg mm^{-2}, whereas diamond has a Knoop hardness of 7000 kg mm^{-2} [44].

The shear modulus is roughly proportional to the hardness. In the past two decades, it has been used as a guide for theoretical predictions of hard materials. Figure 4(a) plots the calculated shear modulus versus the average κ_L. Materials with a superhard property do not necessarily result in a high κ_L. The shear modulus of cubic and wurtzite BAs is only around 123 GPa which is the lowest among the top-14 materials, but their thermal conductivities exceed 2500 $Wm^{-1}K^{-1}$. Another example is superhard cubic silicon nitride (Si_3N_4). Although its reported Vickers hardness is around 35 GPa [45], the calculated thermal conductivity is only around 81 $Wm^{-1}K^{-1}$ [46], which is much lower than that of $BeCN_2$ with the same order of hardness.

Figures 4(b)–4(d) show the average group velocity, heat capacity, and phonon relaxation time of the top-14 materials, respectively. Cubic and wurtzite BAs have group velocities lower than other hard materials, but they have the highest relaxation times. Consequently, they have a high κ_L comparable with diamond. Although superhardness means large elastic constants and group velocities of phonons in the

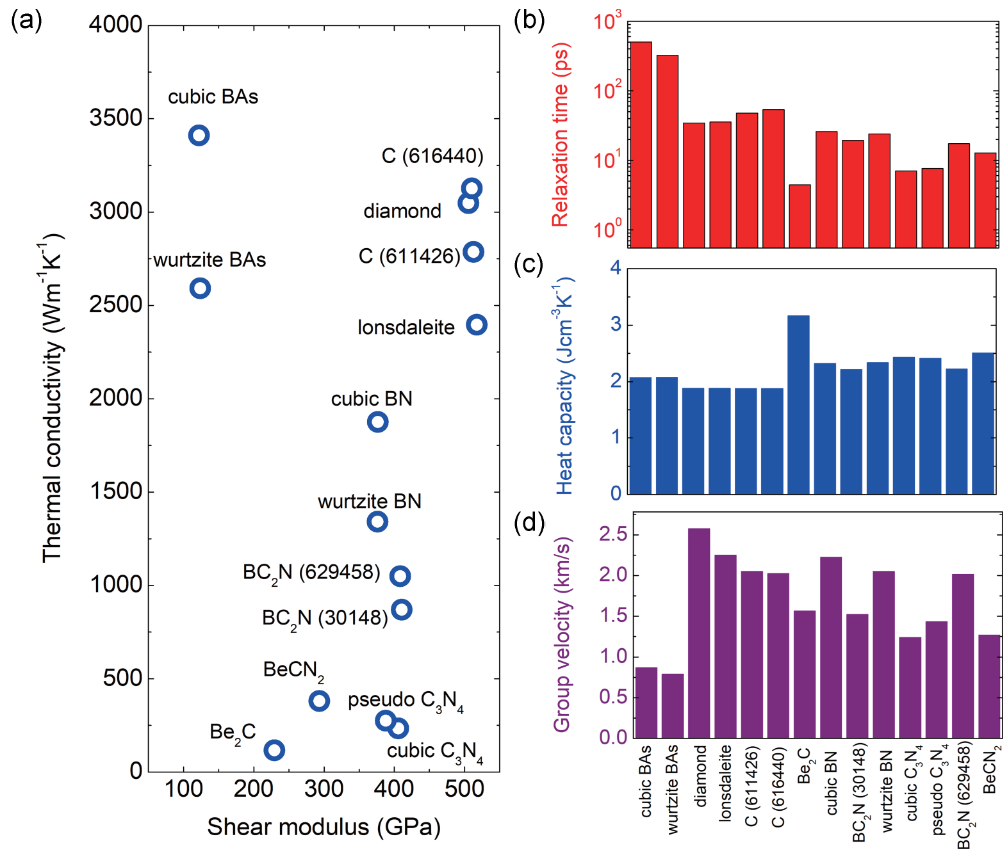

FIG. 4. Hardness versus thermal conductivity and comparison of parameters related with thermal conductivity. (a) Thermal conductivity versus shear modulus for the top-14 materials. (b)–(d) Average group velocity, heat capacity, and relaxation time of the top-14 materials.

linear dispersion regime, κ_L is also determined by other factors such as the phonon relaxation time (i.e., P_3 and anharmonic scattering amplitude). In addition, the properties of phonons with nonlinear dispersions largely influence the κ_L. This highlights the necessity and importance of developing rapid screening models to explore high-κ_L crystals from databases.

Superhard materials consisting of B, C, and N atoms like BN, BC_2N, and C_3N_4 are advantageous over diamond in terms of stability and oxidation because the covalent bond energies between B-N (−117.19 eV) and C-N (−141.74 eV) are stronger than that of C-C (−103.64 eV) [47]. Mixing diamond with BN as a starting material may create new BCN alloy compounds under high pressure and temperature, which are more stable thermally and chemically than diamond and harder than BN.

Among ternary BCN compounds, heterodiamond in the form of BC_2N has gained some attention. Heterodiamond has various structural forms ranging from layered graphitelike and diamond structures. However, the top-14 list includes two cubic BC_2N (mp-ID: 30148 and 629458) structures with thermal conductivities around 784–1392 $Wm^{-1}K^{-1}$. Polycrystalline cubic BC_2N materials have been synthesized from hexagonal BN at 20 GPa and 2200–2250 K [48] and from graphitelike BC_2N above 18 GPa and 2100–2200 K [40,41]. The measured hardness of synthesized cubic BC_2N is higher than that of a cubic BN single crystal but lower than diamond [40,41]. The current finding that cubic BC_2N has a high potential to be an ultrahigh thermal conductor is a motivation to improve the synthesis techniques of heterodiamond.

C_3N_4 is another interesting superhard material with a reported hardness around 37–90 GPa [42]. Here we found two cubic structures of C_3N_4 in the top-14 list. One is pseudo C_3N_4 (mp-ID: 571653) and the other is cubic C_3N_4 (mp-ID: 2852) (Fig. 3). The former has a defect zinc-blende structure with a hole in the central region of the unit cell. The latter has as many as 14 atoms in the primitive unit cell (see Supplemental Material, Fig. S3 [23]), which gives rise to complex phonon modes with many branches. Despite their apparent defective and complex structures, their thermal conductivities exceed 200 $Wm^{-1}K^{-1}$. Efforts have been made to synthesize different phases of carbon nitrides [49]. Martin-Gil *et al.* [49] synthesized a pseudo C_3N_4 by a chemical precursor route under 800 °C and claims the process is scalable.

Lonsdaleite, which is also a wurtzite structure, has been studied previously as its hardness is comparable or even harder than diamond [50]. Its thermal conductivity (2122–2533 $Wm^{-1}K^{-1}$) is also comparable with diamond. Recently, polycrystalline lonsdaleite has been successfully synthesized in a diamond anvil cell at 100 GPa and 400 °C [51] using graphitic layers, which provide a low-energy barrier for progressive transformation from graphite to lonsdaleite. The synthesis temperatures are well below those previously reported for lonsdaleite [50].

The other two hexagonal diamonds (mp-ID: 611426 and 616440) combine the structure features of diamond and

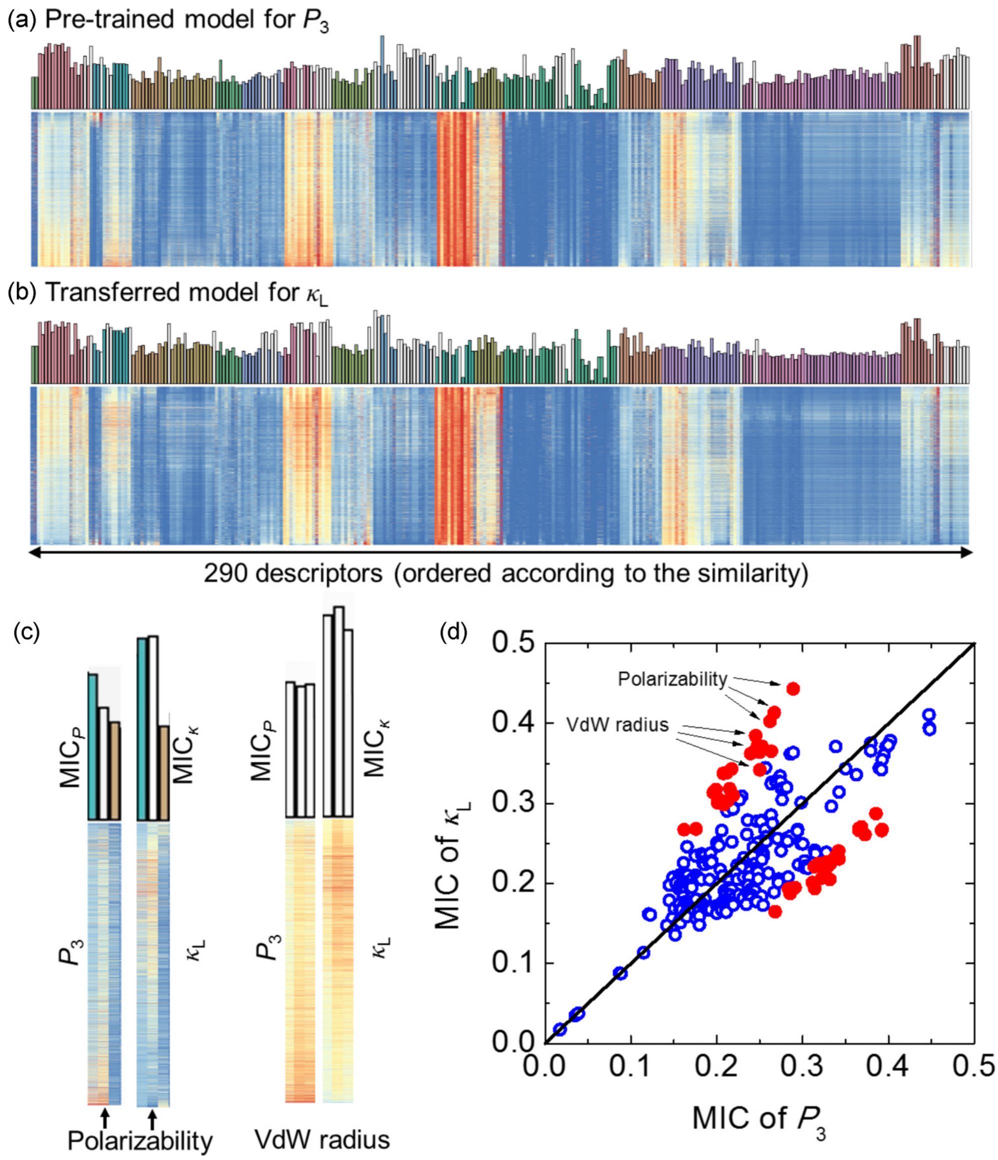

FIG. 5. Descriptor-property heat map and maximal information coefficient (MIC) scores. Descriptor-property heat map of (a) the pretrained model for P_3 and (b) transferred model for κ_L. Top bar plots show the MIC scores for each descriptor. (c) Enlarged heat map for the descriptors of polarizability and the van der Waals (VdW) radius. (d) Distribution of the MIC scores for the 290 descriptors with respect to P_3 and κ_L. Solid red dots denote 57 key descriptors with the MIC scores exhibiting significant differences between P_3 and κ_L (see Supplemental Material, Tables S5 and S6 [23]).

lonsdaleite (see Supplemental Material, Fig. S4 [23]). These can be described as diamond-lonsdaleite superlattices. Their thermal conductivities are comparable with diamond.

D. Knowledge gained from transfer learning

The comparison between the pretrained and transferred models provides some important physical indications. Here, the pretrained model only involves the harmonic phonon properties, whereas the transferred model also requires anharmonic properties (i.e., the magnitude of the three-phonon scattering obtained from cubic interatomic force constants). The success of transfer learning means that the correlation between P_3 and κ_L can be learned from that between the basic crystal structure information and P_3, revealing an underlying commonality of the descriptors corresponding to the harmonic and anharmonic properties.

To understand the differences in how to recognize structure-property relationships for the pretrained and transferred models, we created a descriptor-property heat map (Fig. 5). For each model, the 290 descriptor vectors of the ~60 000 candidates to be screened are displayed onto the heat map. The candidate materials in the heat map are sorted according to a descending order of the predicted values of P_3 or κ_L. This visualization reveals the presence of key descriptors relevant to the pattern recognition inherent in the trained model. Irrelevant or relevant descriptors might exhibit random or nonrandom patterns such as a linear trend along with the ordered predicted properties from top to bottom.

We investigated the underlying mechanisms responsible for the successful transfer from P_3 to κ_L. We aimed to interpret key features that distinguish between the pretrained and post-transferred models. To quantitatively assess the dependency (relevance or association) between each descriptor and the predicted P_3 or κ_L, we introduced the maximal information coefficient (MIC). MIC is a common measure of nonlinear correlations in bivariate random variables [52].

The bar plots in Fig. 5 show the MIC scores of the 290 descriptors with regard to P_3 and κ_L. While most of the realized MICs do not change significantly between P_3 and κ_L, some descriptors exhibit either a significant increase or decrease in the MICs. By looking at the descriptors discriminating the mechanisms regulating P_3 and κ_L where the differences of MIC_P and MIC_κ are ± 0.09 (see Supplemental Material, Tables S5 and S6 [23]), we can identify descriptors that are relevant to the difference between P_3 and κ_L.

The descriptors related with the average or maximum dipole polarizability and van der Waals (vdW) radius over all atoms in the crystal compounds are relevant to κ_L but not to P_3. The polarizability generally correlates with the interactions between electrons and the nucleus. Atoms with a larger number of electrons or atomic radius tend to have a high polarizability. In crystal compounds with a large electronic polarizability, the displacement or force perturbation can be easily transferred and persists over a long range via the orbital electron interaction with the nucleus. A typical example is rocksalt IV-VI materials like PbTe crystal, where the resonant bonding [53] and the corresponding large anharmonic interatomic-force constants are manifested due to the long-range polarization.

The maximum vdW radius that characterizes the nonbonded interactions between atoms also affects the anharmonicity in crystals. The vdW radius is related with the polarizability via the relation $V_\mathrm{w} = \alpha/(4\pi\epsilon_0)$, where α is the polarizability, ϵ_0 is the relative permittivity, and V_w is the vdW volume, which is given by the vdW radius. Since the descriptors correlated with κ_L but not with P_3 should govern the linear output layer between the subnetwork pretrained by P_3 and the final κ_L, it makes sense that polarizability and the vdW radius related to anharmonicity are the descriptors. However, it is meaningful to quantitatively identify that the average or maximum dipole polarizability and the vdW radius are the leading descriptors among those that can also be qualitatively related to anharmonicity. This provides a search direction. In addition to a small P_3, the low average or maximum dipole polarizability and vdW radius are necessary to achieve a high κ_L.

IV. CONCLUSIONS

In summary, we screened over 60 000 crystal compounds with phonon P_3 as the feature quantity and identified a set of semiconducting compounds with high thermal conductivities. Screening was performed based on our developed feature-based transfer learning, which bridges the gaps between the "big data" required for credible machine learning and the "small data" of thermal conductivity. Transfer learning directly models the connection between the basic crystal information and the thermal conductivity with a neural network by transferring descriptors acquired through pretraining for P_3. The successful prediction of high-thermal-conductivity crystals demonstrates the advantage of extrapolative prediction via transfer learning, and reveals the descriptors that are dominantly correlated with the anharmonic phonon properties.

The final, obtained materials in the top-14 list by feature-based and transfer learning screening all show high thermal conductivities, including boron arsenides (BAs), carbon (C), boron nitride (BN), and heterodiamond (BC_2N). They have thermal conductivities on the order of 1000 $\mathrm{W m^{-1} K^{-1}}$, validating the accuracy and high efficiency of the developed screening method. Although most of these are superhard materials with a large group velocity, the results are nontrivial. It has been observed that for some superhard materials, a large dispersion could lead to a large P_3 and limit thermal conductivity. Because screening via P_3 prefers to search crystals with a high thermal conductivity in three different lattice directions, 2D materials such as hexagonal BN and graphite do not appear in our top-14 list due to the weak atomistic interaction in the out-of-plane direction.

The screening also identified known materials that have yet to be studied in the context of heat transport. These include two types of novel carbon crystals with mixed phases of diamond and lonsdaleite and two phases of BC_2N. These materials may be advantageous over well-explored high-thermal-conductivity materials in terms of thermodynamic stability and facileness of synthesis. These findings should contribute to next-generation thermal management technology by broadening the alternatives of high-thermal-conductivity materials and adding degrees of freedom to their surface affinities with various other materials for composite syntheses and integration.

ACKNOWLEDGMENTS

We thank A. Togo for providing the displacement and force data in the phonon database for the P_3 calculation, and T. Shiga for the useful discussion about first-principles calculations. This work was supported by the "Materials research by Information Integration" Initiative ($\mathrm{MI^2I}$) project, CREST Grant No. JPMJCR20Q3, CREST Grant No. JPMJCR19I1, and PRESTO Grant No. JPMJPR16NA from the Japan Science and Technology Agency (JST), and KAKENHI Grants No. 19H00744, No. 19K14902, No. 18K18017, and No. JP17K17762 from the Japan Society for the Promotion of Science (JSPS). The calculations in this work were performed using the supercomputer facilities from JAIST, and the Institute for Solid State Physics, the University of Tokyo.

[1] L. Wei, P. K. Kuo, R. L. Thomas, T. R. Anthony, and W. F. Banholzer, Phys. Rev. Lett. **70**, 3764 (1993).

[2] C. Q. Sun, H. Xie, W. Zhang, H. Ye, and P. Hing, J. Phys. D: Appl. Phys. **33**, 2196 (2000).

[3] J.-K. Lee, M. W. Anderson, F. A. Gray, and P. John, Diam. Relat. Mater. **15**, 1206 (2006).

[4] Q. Cai, D. Scullion, W. Gan, A. Falin, S. Zhang, K. Watanabe, T. Taniguchi, Y. Chen, E. J. G. Santos, and L. H. Li, Sci. Adv. **5**, eaav0129 (2019).

[5] E. Langenberg, E. Ferreiro-Vila, V. Leborán, A. O. Fumega, V. Pardo, and F. Rivadulla, APL Mater. **4**, 104815 (2016).

[6] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, APL Mater. **1**, 011002 (2013).

[7] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, Comput. Mater. Sci. **58**, 227 (2012).

[8] A. Belsky, M. Hellenbrandt, V. L. Karen, and P. Luksch, Acta Crystallogr., Sect. B: Struct. Sci. **58**, 364 (2002).

[9] Y. Xu, M. Yamazaki, and P. Villars, Jpn. J. Appl. Phys. **50**, 11RH02 (2011).

[10] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nat. Mater. **12**, 191 (2013).

[11] A. van Roekeghem, J. Carrete, C. Oses, S. Curtarolo, and N. Mingo, Phys. Rev. X **6**, 041061 (2016).

[12] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Phys. Rev. Lett. **115**, 205901 (2015).

[13] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, Phys. Rev. X **4**, 011019 (2014).

[14] L. Himanen, A. Geurts, A. S. Foster, and P. Rinke, Adv. Sci. **6**, 1900808 (2019).

[15] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, npj Comput. Mater. **5**, 83 (2019).

[16] D. J. Audus and J. J. de Pablo, ACS Macro Lett. **6**, 1078 (2017).

[17] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, in *Advances in Neural Information Processing Systems*, edited by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger (Curran Associates, Red Hook, NY, 2014), Vol. 27.

[18] M. L. Hutchinson, E. Antono, B. M. Gibbons, S. Paradiso, J. Ling, and B. Meredig, arXiv:1711.05099.

[19] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, A survey on deep transfer learning, in *Artificial Neural Networks and Machine Learning – ICANN 2018*, Lecture Notes in Computer Science, edited by V. Krková, Y. Manolopoulos, B. Hammer, L. Iliadis, and I. Maglogiannis, Vol. 11141 (Springer, Cham, 2018).

[20] K. Weiss, T. M. Khoshgoftaar, and D. Wang, J. Big Data **3**, 9 (2016).

[21] H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, and R. Yoshida, ACS Central Sci. **5**, 1717 (2019).

[22] K. Esfarjani and H. T. Stokes, Phys. Rev. B **77**, 144112 (2008).

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevMaterials.5.053801 for lists of training and prediction data, descriptors, and methodology details.

[24] http://phonondb.mtl.kyoto-u.ac.jp/.

[25] T. Tadano, Y. Gohda, and S. Tsuneyuki, J. Phys.: Condens. Matter **26**, 225402 (2014).

[26] https://github.com/yoshida-lab/XenonPy.

[27] L. Lindsay, D. A. Broido, and T. L. Reinecke, Phys. Rev. Lett. **111**, 025901 (2013).

[28] F. Tian, B. Song, B. Lv, J. Sun, S. Huyan, Q. Wu, J. Mao, Y. Ni, Z. Ding, S. Huberman, T.-H. Liu, G. Chen, S. Chen, C.-W. Chu, and Z. Ren, Appl. Phys. Lett. **112**, 031903 (2018).

[29] A. Togo, L. Chaput, and I. Tanaka, Phys. Rev. B **91**, 094306 (2015).

[30] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, Comput. Phys. Commun. **185**, 1747 (2014).

[31] N. Novikov, T. Osetinskaya, A. Shul'zhenko, A. Podoba, A. Sokolov, and I. Petmsha, Dopov. Akad. Nauk. Ukr. RSR, Ser. A Fiz. Mat. Tekh. Nauki USSR, 72 (1983).

[32] J. Kim, D. A. Evans, D. P. Sellan, O. M. Williams, E. Ou, A. H. Cowley, and L. Shi, Appl. Phys. Lett. **108**, 201905 (2016).

[33] B. Lv, Y. Lan, X. Wang, Q. Zhang, Y. Hu, A. J. Jacobson, D. Broido, G. Chen, Z. Ren, and C.-W. Chu, Appl. Phys. Lett. **106**, 074105 (2015).

[34] T. Feng, L. Lindsay, and X. Ruan, Phys. Rev. B **96**, 161201(R) (2017).

[35] S. Li, Q. Zheng, Y. Lv, X. Liu, X. Wang, P. Y. Huang, D. G. Cahill, and B. Lv, Science **361**, 579 (2018).

[36] J. S. Kang, M. Li, H. Wu, H. Nguyen, and Y. Hu, Science **361**, 575 (2018).

[37] F. Tian, B. Song, X. Chen, N. K. Ravichandran, Y. Lv, K. Chen, S. Sullivan, J. Kim, Y. Zhou, T.-H. Liu, M. Goni, Z. Ding, J. Sun, G. A. G. Udalamatta Gamage, H. Sun, H. Ziyaee, S. Huyan, L. Deng, J. Zhou, A. J. Schmidt *et al.*, Science **361**, 582 (2018).

[38] L. Vel, G. Demazeau, and J. Etourneau, Mater. Sci. Eng., B **10**, 149 (1991).

[39] M. Deura, K. Kutsukake, Y. Ohno, I. Yonenaga, and T. Taniguchi, Jpn. J. Appl. Phys. **56**, 030301 (2017).

[40] V. L. Solozhenko, D. Andrault, G. Fiquet, M. Mezouar, and D. C. Rubie, Appl. Phys. Lett. **78**, 1385 (2001).

[41] V. L. Solozhenko, S. N. Dub, and N. V. Novikov, Diam. Relat. Mater. **10**, 2228 (2001).

[42] Q. Fan, C. Chai, Q. Wei, and Y. Yang, Materials (Basel, Switzerland) **9**, 427 (2016).

[43] H.-Y. Gou, F.-M. Gao, J.-W. Zhang, and Z.-P. Li, Chin. Phys. B **20**, 016201 (2011).

[44] G. A. Somorjai and Y. Li, *Introduction to Surface Chemistry and Catalysis,* 2nd ed. (Wiley, Hoboken, NJ, 2010).

[45] N. Nishiyama, R. Ishikawa, H. Ohfuji, H. Marquardt, A. Kurnosov, T. Taniguchi, B.-N. Kim, H. Yoshida, A. Masuno, J. Bednarcik, E. Kulik, Y. Ikuhara, F. Wakai, and T. Irifune, Sci. Rep. **7**, 44755 (2017).

[46] K. Tatsumi, A. Togo, and I. Tanaka, arXiv:1612.08480.

[47] P. Chen, *Material Science and Engineering: Proceedings of the 3rd Annual 2015 International Conference on Material Science and Engineering (ICMSE2015, Guangzhou, Guangdong, China, 15-17 May 2015)* (CRC Press, Boca Raton, FL, 2016).

[48] Y. Zhao, D. W. He, L. L. Daemen, T. D. Shen, R. B. Schwarz, Y. Zhu, D. L. Bish, J. Huang, J. Zhang, G. Shen *et al.*, J. Mater. Res. **17**, 3139 (2002).

[49] J. Martin-Gil, F. J. Martin-Gil, M. Sarikaya, M. Qian, M. José-Yacamán, and A. Rubio, J. Appl. Phys. **81**, 2555 (1997).

[50] F. P. Bundy and J. S. Kasper, J. Chem. Phys. **46**, 3437 (1967).

[51] T. B. Shiell, D. G. McCulloch, J. E. Bradby, B. Haberl, R. Boehler, and D. R. McKenzie, Sci. Rep. **6**, 37232 (2016).

[52] T. Speed, Science **334**, 1502 (2011).

[53] S. Lee, K. Esfarjani, T. Luo, J. Zhou, Z. Tian, and G. Chen, Nat. Commun. **5**, 3525 (2014).