Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling

Jesús Carrete, Wu Li, and Natalio Mingo*
CEA-Grenoble, 17 Rue des Martyrs, Grenoble 38054, France

Shidong Wang, and Stefano Curtarolo†
Center for Materials Genomics, Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, North Carolina 27708, USA

(Received 26 August 2013; revised manuscript received 11 December 2013; published 19 February 2014)

The lattice thermal conductivity (κω) is a key property for many potential applications of compounds. Discovery of materials with very low or high κω remains an experimental challenge due to high costs and time-consuming synthesis procedures. High-throughput computational prescreening is a valuable approach for significantly reducing the set of candidate compounds. In this article, we introduce efficient methods for reliably estimating the bulk κω for a large number of compounds. The algorithms are based on a combination of machine-learning algorithms, physical insights, and automatic ab initio calculations. We scanned approximately 79,000 half-Heusler entries in the AFLOWLIB.org database. Among the 450 mechanically stable ordered semiconductors identified, we find that κω spans more than 2 orders of magnitude—a much larger range than that previously thought. κω is lowest for compounds whose elements in equivalent positions have large atomic radii. We then perform a thorough screening of thermodynamical stability that allows us to reduce the list to 75 systems. We then provide a quantitative estimate of κω for this selected range of systems. Three semiconductors having κω < 5 Wm⁻¹K⁻¹ are proposed for further experimental study.

DOI: 10.1103/PhysRevX.4.011019 Subject Areas: Computational Physics, Condensed Matter Physics, Materials Science

I. INTRODUCTION

High-throughput (HT) computational materials science is a rapidly expanding area of materials research. It merges a plethora of techniques from a variety of disciplines. These include the kinetics and thermodynamics of materials, solid-state physics, artificial intelligence, computer science, and statistics [1]. The application of HT has recently led to new insights and novel compounds in different fields [2–9]. Despite the importance of thermal transport properties for many crucial technologies, there are to date no high-throughput investigations into lattice thermal conductivity.

Here, we seek to address this challenge. We concentrate on the lattice thermal conductivity of half-Heusler (HH) compounds, as they have great promise for applications as thermoelectric materials [10–13]. Half-Heusler compounds are ternary solids. Their crystalline structure consists of two atoms (A and B), located in equivalent positions in a rock-salt structure. A third atom (X) sits in an inequivalent position, filling half of the octahedrally coordinated sites [Fig. 1a].

Experimental studies have reported the thermoelectric figure of merit for a small set of these systems and their alloys [14–18]. Theoretical electronic characterizations have been performed for 36 candidates [19]. It has been speculated that their high thermal conductivity, close to 10 Wm⁻¹K⁻¹, could limit thermoelectric performance [20,21]. At room temperature, the lattice thermal conductivity κω represents the largest contribution to the total conductivity.

Promising thermoelectric figures of merit have been reported both for n-type (1.5 at 700 K [22]) and for p-type (0.8 at 1000 K [17]) half-Heuslers. Such values are comparable to the best thermoelectric materials proposed thus far [23]. Those values, however, were not found in ordered half-Heuslers but rather in alloyed or nanostructured systems. Furthermore, finding ordered compounds with very low κω is advantageous, as their electronic mobilities are expected to be higher than in alloys. In addition, alloying the already low-κω ordered compounds would lower κω even further.

The pool of candidate compounds analyzed in this article is larger than in previous investigations. All possible half-Heusler compounds from all combinations of nonradioactive elements in the periodic table are considered, as
This allows us to provide estimates of the thermal conductivity and has a much lower computational cost. Our accuracy is better than 50% of the exact values.

We implement a novel approach to compute the lattice thermal conductivity. For all the compounds in the Inorganic Crystal Structure Database (ICSD) [27] and AFLOWLIB.org consortium repository [24], this study.

For the 450 resulting stable half-Heuslers, we compute a set of structural, electronic, and harmonic properties. In principle, one could directly compute the thermal conductivity for all the compounds. The computational requirements for this approach would be prohibitive. To solve this issue, our strategy is to obtain a smaller subset of systems. We use physical insights and machine learning techniques to predict the remaining values. Cross validation shows that the approach is reliable for rapidly identifying low-thermal conductivity compounds.

Once the main factors correlated with a low thermal conductivity are identified for the 450-HH library, we use the thermodynamical information in the AFLOWLIB.org database to test the stability of these HHs against more than 110,000 phases. All competing ternary compounds from the Inorganic Crystal Structure Database (ICSD) [27] and all binaries in that database sharing two elements with each HH are included. The final list of thermodynamically stable HH compounds contains 75 entries. For these, we devise and compare with experiment for 75 thermodynamically stable compounds.

II. PREDICTING BULK LATTICE THERMAL CONDUCTIVITIES

The general expression for the thermal conductivity at temperature $T$ is [28]

$$\kappa_\omega = \frac{1}{k_B T^2 V} \sum_\lambda n_0 (n_0 + 1) |v_{\lambda j}^{(c)}|^2 \hbar^2 \omega_{\lambda j}^2 \tau_j,$$  

where $\lambda$ denotes the phonon branch index $\alpha$ and wave vector $\mathbf{q}$, $k_B$ is Boltzmann’s constant, $n_0$ is the Bose-Einstein distribution, $\omega_{\lambda j}$ is the phonon group velocity in the transport direction $z$, and $\tau_j$ is the relaxation time. The relaxation time is determined by third-order derivatives of the total energy with respect to the atomic displacements of any three atoms $i$, $j$, and $k$ in directions $a$, $b$, and $c$ ($\Phi_{\lambda j k}^{abc}$, the anharmonic force constants) in a large supercell [29].

In ordered half-Heuslers, the dominant source of scattering is due to three-phonon processes, and we can calculate thermal conductivities with the full ab initio anharmonic characterization [30,31]. For CoSbZr, one of the most thoroughly studied half-Heuslers, we obtain $\kappa_\omega = 25.0$ Wm$^{-1}$K$^{-1}$. The value is very close to a previous theoretical estimate ($\sim$22 Wm$^{-1}$K$^{-1}$) for monocristalline CoSbZr [19] and slightly higher than the experimental values [32,33] (between 15 and 20 Wm$^{-1}$K$^{-1}$). Synthesized samples of Refs. [32,33] may contain microstructures and imperfections not considered in our work. Despite their accuracy, “full ab initio calculations” of $\kappa_\omega$ [Eq. (1)] are prohibitive for HT studies because of the computational requirements of the derivatives giving $\tau_j$.

In this section, we present two different approaches circumventing the limitation. The first method is based on the empirical observation that the force constants show a high degree of transferability between compounds sharing crystal structure [34]. This suggests that a single set of anharmonic force constants could be used to get an estimate of the bulk $\kappa_\omega$. We call this thermal conductivity calculated with “transferred” forces $\kappa_{\text{trans}}$ (see Table II).

We want to preserve the choice between equivalent positions for maximizing transferability. Thus, instead of

FIG. 1 (color online). (a) Prototype Half-Heusler structure with primitive vectors and a conventional cell. (b) Elements considered in this study.
taking the anharmonic force constants of a particular half-Heusler compound, we choose those of Mg$_2$Si. This compound shares the half-Heusler lattice with sites A and B occupied by Mg atoms. For cross validation, we also fully compute the anharmonic force constants of 32 half-Heusler systems. These are randomly selected with uniform probability inside the convex hull of Fig. 2a, to ensure a wide variety of harmonic and anharmonic features. Comparison between $\kappa_{\omega}$ and $\kappa_{\text{transf}}$ indicates that, although the latter has limited quantitative precision, the qualitative agreement is very good, with a Spearman rank correlation coefficient of 0.93. Hence, the descriptor can be effectively used to separate compounds having high or low $\kappa_{\omega}$. Note that we chose the Spearman rank correlation [35] instead of the usual Pearson one. The former is invariant under any monotonic transformation of one or both variables and takes values ±1 for any strict monotonic (not just linear) dependence.

The second proposed approach is based on a completely different direction: We use “random-forest regression” by leveraging the 32 fully calculated $\kappa_\omega$ as a training set. We can then employ the fitted model to predict the remaining conductivities. We call these predictions $\kappa_{\text{forest}}$ (see Table II). Random forests [36] are a family of general classification and regression algorithms and are well adapted to dependent input data. They have already been successfully applied to numerous problems [37,38], including compound classification [39]. Here, the 32 compounds represent only around 7% of the mechanically stable half-Heuslers. Our input data comprise a large set of descriptor variables, which are expected to correlate with $\kappa_{\omega}$ (Supplemental Material [40]) but is less expensive to obtain. Descriptors include the following:

(i) A priori chemical information: atomic number and weight, position in the periodic table, atomic radius, Pauling electronegativity [41], and Pettifor’s chemical scale $\chi$ (Ref. [42]).

(ii) General compound information: lattice constant $a_{\text{latt}}$, band gap, formation enthalpy, effective masses of electrons and holes, Born effective charges, and dielectric tensor.

(iii) Specific thermal conductivity information: specific heat $c_v$, spherically averaged speed of sound $c_s$, scaled nanograined-limit thermal conductivity $\tilde{\kappa}_{\text{grain}}$, and phase-space volume available for three-phonon scattering processes $P_3$.

After an exploratory phase, we conclude that a satisfactory fit can be safely achieved using only a priori data.

The random-forest method is performed in three steps. First, a large ensemble of decision trees is built by randomly selecting subsets of descriptors and observations. Second, the predictions of all trees are obtained for each data point. Third, the mode (for classification) or the mean (for regression) are taken as the result from the whole ensemble. The algorithm also provides an intrinsic metric to evaluate the importance of each descriptor. This is defined in relation to the effect of randomly permuting the values of that variable on the result [36] (the less resilient upon permutation, the more important).

The prediction of each tree in a random forest can only be a value from the training set, and thus the result of the regression is a weighted average. This average is bounded by the minimum and maximum values within the training data. A small set is unlikely to contain elements having extreme values. Hence, our random-forest regression is expected to have a marked centralizing effect, yielding values tightly grouped around their mean. The frequency densities of both $\kappa_{\text{transf}}$ and this new $\kappa_{\text{forest}}$ are displayed in Fig. 2b. The latter avoids extreme predictions with non-physical magnitudes, a result of the aforementioned centralizing effect.

![FIG. 2 (color online).](image-url)

(a) Joint scatter plot of $c_v$ at 300 K and $P_3$, colored according to our low- and high-$\kappa_{\omega}$ classification based on $\kappa_{\text{forest}}$ (see text); the convex hull of the point set is also included for guidance. (b) Frequency densities of the estimators of thermal conductivity at 300 K $\kappa_{\text{transf}}$ and $\kappa_{\text{forest}}$ as defined in the text. (c) “Violin plot” showing the distribution of $a_{\text{latt}}$ within the low- and high-$\kappa_{\omega}$ classes.
In this sense, machine-learning algorithms outperform crude extrapolations such as those behind $\kappa_{\text{transf}}$. Additionally, $\kappa_{\text{forest}}$ has the advantage that its predictions can be refined with controlled accuracy by changing the size of the training set. Even so, the Spearman rank correlation coefficient between $\kappa_{\text{transf}}$ and $\kappa_{\text{forest}}$ is still 0.66, corroborating the validity of the analysis based on $\kappa_{\text{forest}}$. Furthermore, we find that $\kappa_{\text{forest}}$ is strongly correlated with physical descriptors like $c_a$, $\kappa_{\text{grain}}$, and $P_3$. This confirms our earlier speculation about these methods.

An important concern when training a machine-learning model is whether the training set is diverse or representative enough to justify extrapolating the model to the remaining elements. The values of $\kappa_a$ needed for direct validation of the predicted $\kappa_{\text{forest}}$ are unavailable. Thus, we resort to a repeated fourfold cross validation among the data points in the training set to obtain an estimate of the out-of-sample error. More specifically, we evenly split our training set into four subsets. Then, we obtain a random-forest prediction for the HHs in each of the subsets by using only the remaining 75% of compounds as the new training set. We repeat the process 10 times for different divisions of the data and compute the standard deviation of these predictions. The results are included in Table. I. These estimates support the notion that the model behind $\kappa_{\text{forest}}$ is reasonably insensitive to our choice of training sets. For each cross validation, we compute the Spearman rank correlation coefficient between the out-of-sample random-forest results for the 32 training compounds and their $\kappa_a$. The median value of these Spearman rank correlation coefficients is 0.74, corroborating $\kappa_{\text{forest}}$ as a reliable tool for predicting compound ordering.

The ordering predicted by descriptor $\kappa_{\text{forest}}$ is strongly correlated with that of $\kappa_a$. This allows us to pinpoint the main factors determining high or low thermal conductivities. The bimodal shape of the distribution in Fig. 2b suggests that two groups of half-Heuslers can be identified, with thermal conductivities spread around two different values. A robust version of the “$k$-means” algorithm [43] is employed to optimally place the medians of the low- and high-thermal-conductivity classes at 450 and 23.1 W m$^{-1}$ K$^{-1}$, respectively. By analyzing the importance of variables in the classification, we identify a low Pettifor scale $\chi_X$ and a large average Pauling electronegativity $\bar{e}_{AB}$ as the most critical descriptors for low conductivity (Supplemental Material [40]).

Given the underlying correlations, many different choices can be used for the classification. A trend can even be suggested on the grounds of atomic radii by following a chain of correlations: If the two elements in equivalent positions are chosen so that their average radius is larger than 150 pm, then the probability of the compound being in the low-$\kappa_a$ class is 84%. Physically, this follows from the fact that $\kappa_a$ is highly correlated with the specific heat $c_v$ [Fig. 2a]. The latter is strongly negatively correlated with the lattice parameter $a_{\text{latt}}$. The larger $a_{\text{latt}}$, the lower $c_v$ [44].

In addition, $a_{\text{latt}}$ correlates well with the sum of the atomic radii of the three elements, quantities known $a \text{ priori}$. The atomic radii of the species in positions $X$ concentrate around the average value. This leads to an accurate prediction of $a_{\text{latt}}$ by using only the average atomic radius of the species in positions $X$. This is satisfactory for the screening of materials, as we consider only $X$-type compounds. However, the effect of variations in the $X$ elements on the thermal conductivity is to be considered in the screening of specific families of compounds.

### Table I. Fully calculated thermal conductivities $\kappa_a$ for 32 compounds. These results are then used as the training set for the random-forest predictions. An estimate of the relative standard deviation of $\kappa_{\text{forest}}$ for each compound in the training set, as obtained using repeated fourfold cross validation, is also included. Compounds are always labeled with the element in position $X$ first.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\kappa_a$ (W m$^{-1}$ K$^{-1}$)</th>
<th>$\sigma_{\text{forest},CV}$ (%)</th>
<th></th>
<th>Compound</th>
<th>$\kappa_a$ (W m$^{-1}$ K$^{-1}$)</th>
<th>$\sigma_{\text{forest},CV}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AgKTe</td>
<td>0.508</td>
<td>42</td>
<td></td>
<td>PtGaTa</td>
<td>32.9</td>
<td>11</td>
</tr>
<tr>
<td>BeNaP</td>
<td>4.08</td>
<td>20</td>
<td></td>
<td>PtGeTi</td>
<td>16.9</td>
<td>9.0</td>
</tr>
<tr>
<td>BiBaK</td>
<td>2.19</td>
<td>11</td>
<td></td>
<td>PtLnNb</td>
<td>16.5</td>
<td>8.2</td>
</tr>
<tr>
<td>BiKSr</td>
<td>1.96</td>
<td>6.4</td>
<td></td>
<td>RhHfSb</td>
<td>21.8</td>
<td>13</td>
</tr>
<tr>
<td>BiLiSr</td>
<td>3.04</td>
<td>10</td>
<td></td>
<td>RhNiSi</td>
<td>15.3</td>
<td>11</td>
</tr>
<tr>
<td>CoAsZr</td>
<td>24.0</td>
<td>7.4</td>
<td></td>
<td>RaAsV</td>
<td>23.5</td>
<td>13</td>
</tr>
<tr>
<td>CoBiHf</td>
<td>18.6</td>
<td>14</td>
<td></td>
<td>SbCaK</td>
<td>2.70</td>
<td>9.3</td>
</tr>
<tr>
<td>CoNbZr</td>
<td>25.0</td>
<td>2.4</td>
<td></td>
<td>SiCdSr</td>
<td>13.5</td>
<td>19</td>
</tr>
<tr>
<td>CoSeCo</td>
<td>15.0</td>
<td>13</td>
<td></td>
<td>SnBaSr</td>
<td>2.01</td>
<td>43</td>
</tr>
<tr>
<td>CoSiTa</td>
<td>37.8</td>
<td>7.7</td>
<td></td>
<td>TeAgLi</td>
<td>1.52</td>
<td>11</td>
</tr>
<tr>
<td>FeNbP</td>
<td>109</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table II. Notation for thermal conductivities.

<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_a$</td>
<td>Lattice contribution to $\kappa$ from the “full calculation”</td>
</tr>
<tr>
<td>$\kappa_{\text{transf}}$</td>
<td>Approximated $\kappa_a$ with anharmonic force constants from Mg$_2$Si</td>
</tr>
<tr>
<td>$\kappa_{\text{forest}}$</td>
<td>$\kappa_a$ obtained random-forest regression</td>
</tr>
<tr>
<td>$\kappa_{\text{anh}}$</td>
<td>$\kappa_a$ obtained with four exact anharmonic force constants and a linear model for the rest</td>
</tr>
<tr>
<td>$\kappa_e$</td>
<td>Electronic contribution to $\kappa$</td>
</tr>
<tr>
<td>$\kappa_{\text{grain}}$</td>
<td>Scaled nanograin limit $\kappa_a$</td>
</tr>
</tbody>
</table>
The ingredients of $\kappa_\omega$, for bulk ordered semiconductors depend only on a semilocal characterization of the potential energy surface around the equilibrium configuration. Hence, mechanical stability is sufficient to permit the calculation of the lattice thermal conductivity of a HH. For the analysis performed in the previous section, having the set of 450 mechanically stable HHs reduced and biased by external considerations such as thermodynamical stability would be detrimental to the performance of machine-learning techniques.

On the other hand, in order to propose particular candidates for experimentation, we must maximize the probability that they can be obtained in the laboratory. To this end, we obtain the ternary phase diagrams for each of the 450 mechanically stable HHs. This involves taking into account the formation enthalpies of a large number of possible competing phases. These include but are not limited to all relevant binary and ternary compounds in the ICSD [27]. More specifically, all the elemental compounds, 109,36 binary structures, and 4,363 ternary phases were considered. Many of these phases were already present in AFLOWLIB.org; others were computed specifically for this work. The total number of DFT calculations necessary to obtain the results presented here exceeds 300,000. Our thermodynamic analysis reveals that 77 of the 450 HHs are thermodynamically stable. Spin-polarized calculations reveal that two of the 77 have semimetallic ground states. Then, only the remaining 75 compounds are further considered. The ternary phase diagrams of the final 75 systems are included in the Supplemental Material [40].

Interestingly, all of the 75 predicted stable compounds satisfy the octet or expanded octet rules by virtue of having 8 or 18 valence electrons per unit cell, respectively. We compare these numbers with the frequency distribution of valence electron counts in the initial 79,057-HH library. We conclude that the conditional probabilities of compounds having 8 or 18 valence electrons per unit cell being stable are 1.2% and 3.8%, respectively. While still small, the conditional likelihood of a compound satisfying one of these rules making it through all the filtering steps is much higher than the 0.1% \emph{a priori} probability. Figure 3 shows the distribution of the valence during the reduction of the prototypes’ list.

Even among the reduced list, $\kappa_{\text{forest}}$ still spans more than 1 order of magnitude, its extreme values being 2.33 and 40.3 Wm$^{-1}$K, reinforcing our previous conclusions.

IV. A DESCRIPTOR WITH QUANTITATIVE POWER

Neither of the two descriptors of $\kappa_\omega$ presented so far contains any information about the anharmonic interatomic force constants (IFCs) of each compound. On one hand, the last round of thermodynamical screening puts the number of surviving HHs within the limits of what can be realistically considered for anharmonic calculations. On the other, the qualitative success of $\kappa_{\text{transf}}$ shows that a detailed anharmonic description is not required. To enhance our estimates of the thermal conductivity of stable half-Heuslers, in this section we present a new machine-learning descriptor of $\kappa_\omega$ that integrates only the crucial pieces of the anharmonic properties of the solid. This aids in achieving quantitative accuracy with a much lower computational cost than the full calculation.
Crystallographic symmetries and the equality of mixed partials impose linear constraints on the anharmonic IFCs. With the parameters described in the “Methods” section below and those constraints, we are left with 737 independent anharmonic IFCs per compound. However, many elements of this set are correlated among them, and others are too small to have a decisive role in the value of $\kappa_\omega$. To quantify these assertions, we perform a principal component analysis [48] on the third-order IFCs for the 32 compounds in Table I.

We find that the first four components account for $\sim 99\%$ of the variance in the set. From the results, we can extract an expression for each of the 737 IFCs as a linear combination of those components. Then we perform a multivariable multiple linear regression of the four components on four large and weakly correlated IFCs. By combining the two results, we arrive at a linear model for the whole set of anharmonic IFCs in terms of four parameters that can be obtained with 16 DFT calculations per compound. We use the term $\kappa_\omega$ to describe the third-order IFCs thus reconstructed, and $\kappa_{\text{anh}}$ for the second-order IFCs for each compound.

The blue circles in Fig. 4a show a comparison between $\kappa_{\text{anh}}$ and the exact $\kappa_\omega$ for the 32 compounds in the training set. With two exceptions (compounds with comparatively very high thermal conductivities), this new descriptor yields excellent quantitative estimates of $\kappa_\omega$. Moreover, fourfold cross validation shows that it is insensitive to the particular choice of training set. As a final test, we perform full thermal conductivity calculations for four compounds selected at random from those outside the training set: AgBaSb, AgNaTe, InCdY, and TlLaMg. The results are depicted as red crosses in Fig. 4a. This shows that the quality of the prediction is as good as for the 32 training compounds.

The distribution of $\kappa_{\text{anh}}$ over the 75 thermodynamically stable HHs [Fig. 4b] confirms the presence in the sample of

<table>
<thead>
<tr>
<th>$\kappa_{\text{anh}}$ (Wm$^{-1}$ K$^{-1}$)</th>
<th>$\kappa_{\text{anh}}$ (Wm$^{-1}$ K$^{-1}$)</th>
<th>$\kappa_{\text{anh}}$ (Wm$^{-1}$ K$^{-1}$)</th>
<th>$\kappa_{\text{anh}}$ (Wm$^{-1}$ K$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AuAlHf</td>
<td>16.7</td>
<td>FeAsNb</td>
<td>47.6</td>
</tr>
<tr>
<td>BLiSi</td>
<td>62.1</td>
<td>FeAsTa</td>
<td>32.9</td>
</tr>
<tr>
<td>BiBaK</td>
<td>1.24</td>
<td>FeGeW</td>
<td>32.8</td>
</tr>
<tr>
<td>CoAsHf</td>
<td>20.0</td>
<td>FeNbSb</td>
<td>29.1</td>
</tr>
<tr>
<td>CoAsTi</td>
<td>37.1</td>
<td>FeSbTa</td>
<td>31.2</td>
</tr>
<tr>
<td>CoAsZr</td>
<td>27.7</td>
<td>FeSbV</td>
<td>24.1</td>
</tr>
<tr>
<td>CoBiHf</td>
<td>22.5</td>
<td>FeTeTi</td>
<td>26.2</td>
</tr>
<tr>
<td>CoBiTi</td>
<td>27.1</td>
<td>GeAlLi</td>
<td>16.5</td>
</tr>
<tr>
<td>CoBiZr</td>
<td>17.8</td>
<td>IrAsTi</td>
<td>30.1</td>
</tr>
<tr>
<td>CoGeNb</td>
<td>36.2</td>
<td>IrAsZr</td>
<td>17.4</td>
</tr>
<tr>
<td>CoGeTa</td>
<td>27.2</td>
<td>IrBiZr</td>
<td>12.8</td>
</tr>
<tr>
<td>CoGeV</td>
<td>29.1</td>
<td>IrGeNb</td>
<td>33.0</td>
</tr>
<tr>
<td>CoHfSb</td>
<td>21.9</td>
<td>IrGeTa</td>
<td>37.2</td>
</tr>
<tr>
<td>CoNbSi</td>
<td>30.1</td>
<td>IrGeV</td>
<td>30.0</td>
</tr>
<tr>
<td>CoNbSn</td>
<td>20.7</td>
<td>IrHfSb</td>
<td>24.7</td>
</tr>
<tr>
<td>CoSbTi</td>
<td>23.3</td>
<td>IrHfSn</td>
<td>19.8</td>
</tr>
<tr>
<td>CoSbZr</td>
<td>24.4</td>
<td>IrSnTa</td>
<td>22.1</td>
</tr>
<tr>
<td>CoSiTa</td>
<td>36.9</td>
<td>NiAsSc</td>
<td>17.5</td>
</tr>
<tr>
<td>CoSnTa</td>
<td>22.7</td>
<td>NiBiSc</td>
<td>14.3</td>
</tr>
<tr>
<td>CoSnV</td>
<td>19.8</td>
<td>NiBiY</td>
<td>10.6</td>
</tr>
</tbody>
</table>
compounds with thermal conductivities much lower than 10–20 Wm⁻¹K⁻¹. This is characteristic of experimentally measured HHs. The values of κᵣ₁ for the 75 stable HHs are listed in Table III. Notably, the subset of 10 thermodynamically stable half-Heuslers for which κᵣ₁ was directly computed already contains BiBaK, with κᵣ₁ = 2.20 Wm⁻¹K⁻¹. Outside of the training sample, the lowest κᵣ₁ values are 1.72, 2.84, and 3.49 for PtLaSb, RhLaTe, and SbNaSr, respectively.

V. CONCLUSIONS

In this article, we have presented three computational methods for estimating the bulk κᵣ₁ of a large library of half-Heusler compounds. We surmount the formidable task of full ab initio characterization. We find that κᵣ₁ is spread over more than 2 orders of magnitude over mechanically stable half-Heuslers. This is a much broader range than that suggested by limited experimental available data. By using a set of descriptors and random-forest regression, we have built and tested an effective classification model. We found that compounds are most likely to have low thermal conductivity if the average atomic radius of the atoms in structural positions A and B is large. This also correlates with large lattice parameters and low specific heat.

Extensive thermodynamical calculations allow one to remove compounds with more stable competing phases from the list. We employ our third method, with better quantitative accuracy and higher computational cost, to perform a finer analysis of the distribution of κᵣ₁ over the reduced library. We conclude that ordered half-Heusler compounds with a κᵣ₁ ≲ 3 Wm⁻¹K⁻¹ value (a factor of 3 below the best scenarios for ordered compounds and comparable to alloyed systems) very likely exist. The results corroborate the competitiveness of machine-learning methods in accelerated material design [1].

VI. METHODS

AFLOWLIB library of half-Heusler systems.—The 79,057 half-Heusler systems are calculated with the high-throughput framework AFLOW [4,25,49,50] based on ab initio calculations of the energies by the VASP software [51] with projector augmented wave (PAW) pseudopotentials [52], and Perdew, Burke, and Ernzerhof exchange-correlation functionals [53]. The AFLOWLIB energies are calculated at zero temperature and pressure, with spin polarization and without zero-point motion or lattice vibrations. All crystal structures are fully relaxed (cell volume and shape and the basis atom coordinates inside the cell). Numerical convergence to about 1 meVatom⁻¹ is ensured by a high-energy cutoff (30% higher than the highest-energy cutoff for the pseudopotentials of the components) and by the dense 6,000 k-points per reciprocal atom Monkhorst-Pack meshes [54].

Interatomic force constants.—3 × 3 × 3 supercells are used in second-order IFC calculations. The Phonopy [55] package is used to generate a minimal set of atomic displacements by harnessing the point and translational symmetries of the crystal structure, and custom software was developed in order to do the same in anharmonic IFC calculations. For those calculations, 4 × 4 × 4 supercells are generated and a cutoff radius of 0.85aₙₙ is imposed on the interactions. The 2 × 2 × 2 and 3 × 3 × 3 Monkhorst-Pack k-point grids are employed, and spin polarization is excluded to improve speed.

Solution of the Boltzmann transport equation.—Our self-consistent iterative approach is described in detail in Ref. [30]. Both three-phonon processes and the natural isotopic distribution of each element are taken into account as sources of scattering. A Gaussian smearing scheme with adaptive breadth [31] is chosen for integrations in the Brillouin zone. When using anharmonic IFCs from Mg₂Si to approximate κᵣ₁ for all materials, the solution to the Boltzmann transport equation failed to converge for five compounds, which are consequently excluded from the associated analysis.

Regression and classification.—The R statistical computing environment [56] is chosen for all statistical analyses. Random-forest models are used as implemented in the “randomForest” package [57]. As a check, all regressions and classifications are repeated using a generalized boosted tree algorithm [58]; in all cases, the results are found to be in good agreement with those afforded by random forests.

ACKNOWLEDGMENTS

The authors thank Professor D. Broido and Dr. L. Lindsay for providing us with a set of an harmonic force constants for Mg₂Si, and Dr. A. Stelling, Dr. O. Levy, Professor S. Sanvito, Professor M. Buongiorno Nardelli, and Professor M. Formari for useful comments. This work is partially supported by the French “Carnot” project SIEVE, by DOD-ONR (N00014-13-1-0635, N00014-11-1-0136, and N00014-09-1-0921) and by Duke University—Center for Materials Genomics.


[44] The decreasing trend can be understood by considering that the specific heat per atom at high temperatures relates to the number of degrees of freedom, through the equipartition theorem. Hence, a $c_v \propto a_{\text{lat}}^3$ dependence should be expected. In fact, the observed trend is sharper because of the differences in Debye temperature among the compounds.