



Norwegian University of Life Sciences
Faculty of Science and Technology

Philosophiae Doctor (PhD)
Thesis 2023:13

Identifying Materials with Low Lattice Thermal Conductivity Using Machine Learning and Computational Modeling

Identifikasjon av materialer med lav termisk
gitterledningsevne ved bruk av maskinl ring
og databeregninger

Rasmus Andr  Tran s

Identifying Materials with Low Lattice Thermal Conductivity Using Machine Learning and Computational Modeling

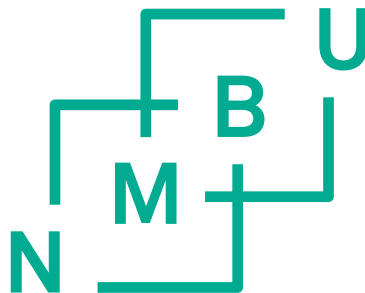
Identifikasjon av materialer med lav termisk gitterledningsevne ved bruk av maskinl ring og databeregninger

Philosophiae Doctor (PhD) Thesis

Rasmus Andr  Tran s

Norwegian University of Life Sciences
Faculty of Science and Technology

 s, 2023



Thesis number 2023:13
ISSN 1894-6402
ISBN 978-82-575-2042-7

Abstract

Lattice thermal conductivity is a key materials property in applications related to thermal functionality, such as thermal barrier coatings, thermal conductors in microelectronics, and solid-state waste-heat recovery devices. The lattice thermal conductivity governs the rate of heat energy transfer in thermoelectric materials, which are materials that can directly convert heat to electricity and vice versa. These materials become interesting in applications that require electricity generation or local cooling. Thermoelectric materials depend on a low lattice thermal conductivity to attain high heat-to-electricity conversion efficiency. The materials used in present thermoelectric generators are often based on toxic or scarce elements. New high-efficiency thermoelectric materials are therefore desired for sustainable and environmentally friendly energy harvesting. Two main research challenges are investigated in this thesis: 1) reducing the lattice thermal conductivity to enhance thermoelectric performance, and 2) identifying new compounds with low lattice thermal conductivity. Addressing these challenges experimentally is a daunting task – especially for 100s or 1000s of compounds – as experiments are costly, time-consuming, and require expert domain knowledge. This thesis, therefore, relies on lattice thermal conductivity from theoretical calculations based on quantum mechanical simulations.

Addressing challenge 1), the lattice thermal conductivity of 122 half-Heusler compounds is calculated using density functional theory and the temperature-dependent effective potential method. Phonon scattering from partial sublattice substitutions and grain boundaries are included in calculations, in an attempt to reduce the lattice thermal conductivity. We find that isovalent substitutions on the site hosting the heaviest atom should be performed to optimally reduce the lattice thermal conductivity in most half-Heuslers. Compounds with large atomic mass differences can have a large drop in lattice thermal conductivity with substitutions. Examples of such compounds are AlSiLi and TiNiPb, which achieve a $\sim 70\%$ reduction of their lattice thermal conductivity when substituting Si by Ge and Pb by Sn at 10 % concentration. The reduction from additional scattering mechanisms enables a handful half-Heuslers to attain a lattice thermal conductivity close to 2 W/Km at 300 K. Calculations for full-Heusler AlVFe₂ reveal that the introduction of 15 % Ru substitutions on the Fe-site and 100 nm grain boundaries can reduce the lattice thermal conductivity from 46 W/Km to 7 W/Km.

Tackling challenge 2) is done by computational screening for low lattice thermal conductivity compounds. Coupling calculations with machine learning accelerates the screening. When training the machine learning model on calculated lattice thermal conductivities, it learns to recognize descriptor patterns for compounds with low lattice thermal conductivity. The size of the training

set is limited by the large computational cost of calculating lattice thermal conductivity. It is therefore challenging to obtain a diverse set of training compounds, especially so because low lattice thermal conductivity compounds tend to be rare. We find that including certain compounds in the training can be crucial for identifying low lattice thermal conductivity compounds. Active sampling enables scouting of the compound space for compounds that should enter the training set. Principal component analysis and Gaussian process regression are used in the active sampling schemes. With Gaussian process regression we screen 1573 cubic compounds, where 34 have predicted lattice thermal conductivity ≤ 1.3 W/Km at 300 K – as well as electronic band gaps – indicating that they could be potential thermoelectric compounds.

The findings in this thesis show that certain compounds could have a drastic reduction in the lattice thermal conductivity with sublattice substitutions. Thermoelectric compounds with favorable electronic properties – but high lattice thermal conductivity – can be investigated in future studies if there is a potential for a large drop in the lattice thermal conductivity with sublattice substitutions. The machine learning and active sampling schemes are scalable, and future works could expand upon this thesis by including different compound classes in training and screening. This would enlarge the search space for promising thermoelectric compounds, increasing the likelihood of encountering high-efficiency candidates. It is also possible to combine the two challenges faced in this thesis. A machine learning model can be trained to predict the lattice thermal conductivity of compounds with sublattice substitutions. This would further increase the pool of possible compounds where promising thermoelectric compounds could reside.

Sammendrag

Termisk gitterledningsevne er en viktig materialeegenskap i tekniske instrumenter som anvender varmeledningsteknologi, slik som termiske barriere-belegg, termiske ledere i mikroelektronikk, og varmegjenvinningsenheter. Denne egenskapen styrer raten av varmeenergi-overføring i termoelektriske materialer. Disse materialene kan omgjøre varmeenergi til elektrisk energi og motsatt, og er derfor lovende i produkter som avhenger av elektrisitetsgenerering eller utnytter lokal kjøling. Termoelektriske materialer må ha lav termisk gitterledningsevne for å opprettholde høy effektivitet. Dagens termoelektriske materialer er ofte basert på giftige eller sjeldne materialer, slik som bly eller tellur. Det er derfor nyttig å finne nye materialer med høy effektivitet for å videre anvende termoelektrisk energi-høsting på en bærekraftig måte. To hovedutfordringer er undersøkt i denne avhandlingen: 1) reduksjon av termisk gitterledningsevne for å øke termoelektrisk effekt, og 2) identifikasjon av nye materialer med lav gitterledningsevne. Å løse disse utfordringene eksperimentelt er krevende siden eksperimenter er dyre, tar mye tid, og krever ekspert-kunnskap. I denne avhandlingen brukes derfor teoretiske beregninger basert på kvantemekaniske simuleringer for å estimere termisk gitterledningsevne.

I arbeidet med utfordring 1) beregnes termisk gitterledningsevne til 122 half-Heusler-materialer basert på temperaturavhengige materialsimuleringer. For å redusere termisk gitterledningsevne inkluderes ekstra fonon-spredningsmekanismer: sub-gitter-substitusjoner (legeringer) og korngrenser. Vi finner at isovalente substitusjoner på gitter-plassen som innehar det tyngste atomet gir den største reduksjonen i termisk gitterledningsevne for de fleste materialene. Materialer med stor atommasse-forskjell kan ha en stor reduksjon i termisk gitterledningsevne med substitusjoner. AlSiLi og TiNiPb er eksempler på slike materialer, og oppnår en $\sim 70\%$ reduksjon i termisk gitterledningsevne når Si er substituert med Ge og Pb er substituert med Sn med 10% konsentrasjon. Reduksjonen fra ekstra spredningsmekanismer gjør at en håndfull half-Heuslere oppnår termisk gitterledningsevne nærme 2 W/Km . Beregninger for full-Heusleren AlVFe₂ viser at introduksjonen av 15% Ru-substitusjon på Fe-gitterplassen og 100 nm korngrenser kan redusere termisk gitterledningsevne fra 46 W/Km til 7 W/Km .

Utfordring 2) er utforsket med data-drevne søk for materialer med lav termisk gitterledningsevne. Teoretiske beregninger sammen med maskinlæring gjør det mulig å søke igjennom flere materialer. Maskinlæringsmodellen lærer å gjenkjenne materialer med lav termisk gitterledningsevne ved å forstå likheter mellom egenskapene til materialene. Fordi det er dyrt å beregne termisk gitterledningsevne blir typiske treningssett små. Vi bruker derfor metoder som aktivt velger hvilke materialer som burde inngå i treningssettet. Prinsipal-

komponent-analyse og regresjon basert på Gaussiske prosesser er anvendt for å finne materialer til trening. Ved bruk av regresjonen søker vi blant ~ 1500 materialer, og finner at 34 har prediktert termisk gitterledningsevne under 1.3 W/Km ved 300 K – i tillegg til elektronisk båndgap – som indikerer at de kan være mulige termoelektriske materialer.

Observasjonene i denne avhandlinga viser at spesifikke materialer kan ha en drastisk reduksjon i termisk gitterledningsevne med sub-gitter-substitusjoner. Termoelektriske materialer med gode elektriske egenskaper – men høy termisk gitterledningsevne – kan bli studert i framtidige studier om de har et stort potensial for reduksjon i termisk gitterledningsevne. Maskinlæringsmodellen og metoden som gjør aktivt utvalg av treningssettet kan skaleres, og framtidig arbeid kan ekspandere på denne avhandlingen ved å inkludere andre materialklasser i treningen og søket. Dette vil øke størrelsen til material-rommet og kan øke sannsynligheten for å finne kandidater med høy termoelektrisk effektivitet. Det er også mulig å kombinere erfaringene fra de to utfordringene i denne avhandlinga. En maskinlæringsmodell kan trenes for å predikere termisk gitterledningsevne til materialer med subgitter-substitusjoner. Dette kan videre øke mengden materialer inkludert i søket og nye termoelektriske materialer kan forhåpentligvis bli funnet.

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* at the Norwegian University of Life Sciences, Faculty of Science and Technology. The work in this thesis was done from 2019 to 2023 at the Department of Mechanical Engineering and Technology Management in the Material Theory and Informatics group. Main supervision was given by Associate Professor Kristian Berland. Regular supervision was also given by two co-supervisors, Professor Ole Martin Løvvik at SINTEF Sustainable Energy Technology and the Department of Physics, University of Oslo, and Associate Professor Oliver Tomic at the Department of Data Science, Norwegian University of Life Sciences.

This doctoral work resulted in three published articles and one article to be submitted. The articles are presented at end of this thesis. The doctoral program has involved taking courses counting in total 30 ECTS, which is equivalent to one semester of study. Most of the work was done at the Norwegian University of Life Sciences in Ås. Some work was also done during sporadic visits to the SINTEF office in Oslo. During COVID-19 lockdowns I mostly worked at home office in Trøndelag. March to June 2022 was spent in Vienna, Austria. This research stay was hosted by Professor Georg K. H. Madsen in the Theoretical Materials Chemistry group at the Vienna University of Technology.

Parts of the work in this thesis were inspired by the research project Allotherm (Project No. 314778), funded by the Research Council of Norway. The goal of this project is to use theoretical calculations to identify promising thermoelectric alloy candidates and verify the thermoelectric performance with experiments.

The PhD was internally funded by the Norwegian University of Life Sciences. The Norwegian e-infrastructure for research and education, Sigma2, granted access to the national high-performance computing clusters used for computations in the thesis.

Acknowledgements

First, I would like to thank my main supervisor, Kristian Berland. His supervision enabled me to improve my scientific writing and communication, do research, and attain a new level of attention to detail. Whenever I ran out of creativity or patience I could always rely on Kristian. For taking me on as the group's first PhD candidate, he has my gratitude.

I would like to thank my co-supervisors Ole Martin Løvvik and Oliver Tomic. They taught me scientific methods that were used throughout the work and motivated me to pursue my scientific goals. I am also grateful that Ole Martin introduced me to his research group at SINTEF and that Oliver introduced me to the Data Science Department at the Norwegian University of Life Sciences.

Georg K. H. Madsen has my gratitude for allowing me to visit his research group at TU Wien in the spring of 2022. The group members welcomed me with open arms and made my trip to Austria a highlight of my PhD both scientifically and socially.

I want to thank all group members in the Materials Theory and Informatics Group. All the coffee breaks, chats, lunches, and friendly banter have been welcome and much needed for the completion of this PhD. I want to highlight the coffee chats with Elin and Mojtaba during the COVID-19 lockdowns as well. Although our meetings were digital it was always pleasant to see your faces on the screen. Now you will soon be the senior PhD candidates in the group, I hope you enjoy the experience!

I would like to thank the following people for feedback on the thesis and proof-reading: Elin Dypvik Sødahl, Seyedmojtaba Seyedraoufi, Øven Andreas Grimenes, Sebastian Bichelmaier, and Siri-Unn Sagen.

Lastly, my thanks to my friends, family, and my girlfriend who all supported and encouraged me during my work. Both between and during COVID-19 lockdowns you have shown me that there is an interesting world outside of the PhD universe. I am grateful for all the fine moments we have shared, and I am sure there are many more to come.

Foreword: What is a PhD?

What is a PhD and what is its purpose? These questions have been presented to me on several occasions during the last three years. To present the PhD education I will try to conceptualize it in comparison to what is known by most people outside academia. In Fig. 1 there is a sketch of what we typically regard as "formal" education. When we start school, we acquire a broad set of skills. This set of skills is "common", in that almost anyone (that is, almost anyone within the constellation) can familiarize themselves with the topics and themes studied. In high school we tend to specialize to a small degree. Here, again most knowledge acquired is familiar to most, and some time is spent developing skills such as writing, mathematics, and communication. A B.Sc. is more specialized. It will either prepare the student for a specified occupation or lay the foundation for further studies. An M.Sc. is much specialized. Special courses are taken to build up to a thesis. The M.Sc. thesis is proof that the student has a good understanding of the specialized topic. Now, on to the PhD. The PhD thesis is typically so specialized that often only a handful of people will be able to fully understand its contents. Likewise, as with the M.Sc, the thesis displays that the student has a solid understanding of its topics. An important aspect that differentiates the PhD education from others is that the goal is to discover something new scientifically. Through the PhD one should contribute to ones research field, and present the discoveries in a way such that other researchers can build upon it. This would be the sliver of the blue bulge outside the green disc in Fig. 1, expanding the limits of our knowledge. In essence, it is training to be a researcher while being a researcher, akin to laying the tracks with the train running. The PhD lies in the intersection between being a student of a scientific discipline and a contributor to the discipline. In short, during the works of a PhD, one is a scientific apprentice.

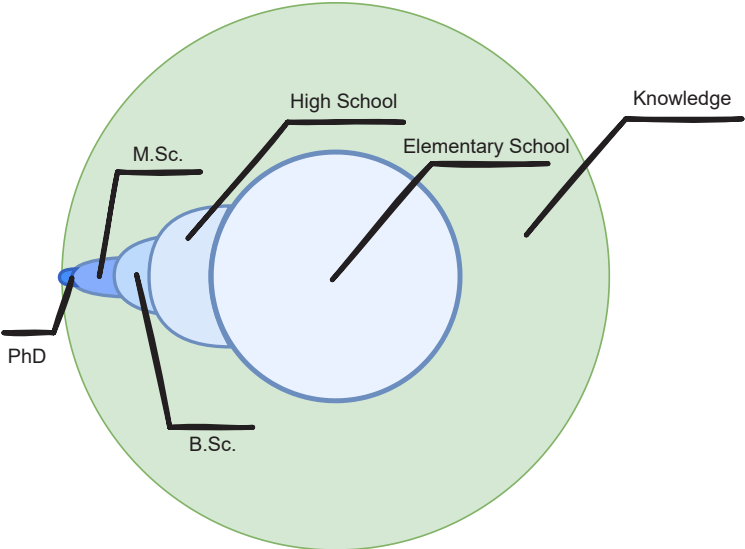


Figure 1: Progress of learning throughout education, adapted from [1].

List of Articles

Article I

R. Tranås, O. M. Løvvik, O. Tomic, and K. Berland. "Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis", *Comput. Materi. Sci*, 202, 2022, 110938. <https://doi.org/10.1016/j.commatsci.2021.110938>.

STATUS: PUBLISHED

This article presents a machine learning model for predicting the lattice thermal conductivity of half-Heusler compounds. It contains the lattice thermal conductivity of 122 compounds, which is calculated using density functional theory and the temperature-dependent effective potential method. An active sampling scheme based on principal component analysis is used to identify which compounds should enter into the training set for identifying low lattice thermal conductivity compounds. I did the lattice thermal conductivity calculations, the machine learning, made the figures, and wrote the draft. Writing and revisions following the draft I did jointly with the co-authors. I presented preliminary results from this article in a talk at the APS March Meeting 2021, and in a poster at Multiscale Modelling of Materials and Molecules 2021.

Article II

R. Tranås, O. M. Løvvik, and K. Berland. "Attaining low lattice thermal conductivity in half-Heusler sublattice solid solutions: which substitution site is most effective?", *Electron. Mater.*, 2022, 3(1), 1-14. <https://doi.org/10.3390/electronicmat3010001>.

STATUS: PUBLISHED

This article concerns an investigation of the effects of mass-disorder and grain-boundary scattering on the lattice thermal conductivity of 122 half-Heusler compounds. The lattice thermal conductivity is calculated similarly to as in Article I. The main findings in the article are that the choice of substitution site can greatly affect the reduction in the lattice thermal conductivity and that it is often beneficial to substitute on the site hosting the heaviest atom in the parent compound. I did the calculations, made the figures, and did most of the data analysis. I wrote the draft, and jointly wrote the manuscript with the co-authors. I presented contents of this article in a poster presentation at the EMRS Spring Meeting 2022.

Article III

R. Tranås, O. M. Løvvik, and K. Berland. "Lattice thermal conductivity of cubic compounds with iterative training over material space using Gaussian process regression"

STATUS: TO BE SUBMITTED

This article presents a study using machine learning to predict the lattice thermal conductivity of 1573 cubic compounds. A Gaussian process regression model is trained using lattice thermal conductivities calculated based on density functional theory and the temperature-dependent effective potential method. The uncertainties from the machine learning model enable active sampling of the cubic compounds. We identify new low lattice thermal conductivity compounds with bandgaps. Two compounds, Na_2TlSb and Ca_3AsBr_3 , are investigated to find the origin of their low lattice thermal conductivity. I did the lattice thermal conductivity calculations, trained the machine learning model, and wrote the code for extracting features. I wrote the draft, made the figures, and jointly wrote the manuscript with the co-authors. I conceptualized the article together with K. Berland. I presented results from this article in a poster presentation at ECT 2022 and Nordic AI Meet 2022.

Article IV

K. Berland, O. M. Løvvik, and R. Tranås. "Discarded gems: Thermoelectric performance of materials with band gap emerging at the hybrid-functional level", *Appl. Phys. Lett.*, 2021, 119(8), 081902. <https://doi.org/10.1063/5.0058685>.

STATUS: PUBLISHED

This article presents a screening study of cubic compounds in the Materials Project database where the goal is to find promising thermoelectric candidates. Electronic band gaps are often underestimated with standard density functionals. Therefore, the band gaps in this work are assessed using a hybrid functional. Eight compounds are found to have no bandgap at standard density functional theory level, but non-zero bandgap using the hybrid functional. Two compounds, MgSc_2Hg and Li_2CaSi , display promising thermoelectric properties. I wrote the code to extract the compounds from the Materials Project and did lattice thermal conductivity calculations for the eight compounds. I contributed with text editing and fixing spelling errors.

Additional Scientific Work

Oral Presentations

- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Building intuition on lattice thermal conductivity in half-Heusler materials with machine learning feature selection. Presented at: *Virtual Conference on Thermoelectrics 2020*, 2020-07-21–2020-07-23
- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Current state of own research on identifying low lattice thermal conductivity materials using quantum mechanics simulations and machine learning. Presented at: *NMBU Alumni scientific presentation*, 2020-11-10–2020-11-10
- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Active sample selection for improving machine learning predictions of low lattice thermal conductivity materials: half-Heusler case study and beyond. Presented at: *ASMOD Presentation of tools*, 2021-04-19–2021-04-19
- O. M. Løvvik, M. Schrade, Ø. A. Grimenes, R. Tranås, and K. Berland. Predictive screening for thermoelectric properties using atomic-scale simulations and machine learning tools. Presented at: *The Leibniz Institute for Solid State and Materials Research*, 2022-05-03–2022-05-03
- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Discovering Energy Materials: Low Thermal Conductivity Thermoelectric Compounds and Barocaloric Compounds. Presented at: *TU Wien, Theoretical Materials Chemistry*, 2022-03-27–2022-03-27

Poster Presentations

- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Fingerprints of low lattice thermal conductivity compounds: half-Heusler case study and beyond. Presented at: *APS March Meeting 2021*, 2021-03-15–2021-03-19
- R. Tranås, O. M., Løvvik, O. Tomic, and K. Berland. Active sample selection for finding low lattice thermal conductivity materials using machine learning. Presented at: *Multiscale Modelling of Materials and Molecules 2021*, 2021-06-07–2021-06-08
- R. Tranås, O. M., Løvvik, and K. Berland. Alloying leads to drastic reduction of lattice thermal conductivity of half-Heusler compounds. Presented at: *EMRS Spring Meeting 2022*, 2022-05-30–2022-06-03

- R. Tranås, G. H. K. Madsen, J. Carrete, and K. Berland. Predictive Modeling of Order-Disorder Phase Transitions in Hybrid Organic Materials with Machine Learning Force Fields. Presented at: *European Crystallography Conference 2022*, 2022-08-23–2022-08-27
- R. Tranås, O. M., Løvvik, and K. Berland. Finding Low Lattice Thermal Conductivity Compounds in Materials Space: Machine Learning with Active Sampling. Presented at: *ECT 2022*, 2022-09-14–2022-09-16
- R. Tranås, O. M., Løvvik, and K. Berland. Optimal alloying site for reducing lattice thermal conductivity of half-Heuslers. Presented at: *ECT 2022*, 2022-09-14–2022-09-16
- R. Tranås, O. M., Løvvik, and K. Berland. Material Informatics: Machine Learning with Active Sampling for Small Training Sets. Presented at: *Nordic AI Meet 2022*, 2022-11-14–2022-11-15

Supervision

- Ø. A. Grimenes, "Screening of thermoelectric performance of half-Heusler materials and their alloys from first principles", Masters Thesis, Norwegian University of Life Sciences, Ås, Norway, 2021 [2].

Contents

Abstract	iii
Sammendrag	v
Preface	vii
Acknowledgements	ix
Foreword: What is a PhD?	xi
List of Articles	xiii
Additional Scientific Work	xv
Contents	xvii
List of Figures	xxi
List of Tables	xxv
1 Introduction	1
1.1 Motivation and Background	1
1.2 Thesis Introduction and Articles	3
1.3 Outline	4
2 Materials and Methods	5
2.1 Thermoelectricity: Principles and Materials	5
2.1.1 Seebeck and Peltier Effect	6
2.1.2 Thermoelectric Figure of Merit	6
2.1.3 Thermoelectric Materials	7
2.1.4 Relevance for the Articles	9
2.2 Lattice Dynamics	11
2.2.1 Lattice Potential Energy and Phonons	11
2.2.2 Phonon Dispersions of Si	13
2.2.3 Lattice Thermal Conductivity	13
2.2.4 Three-Phonon, Mass-Disorder, and Grain- Boundary Scattering	15
2.2.5 Lattice Thermal Conductivity Models	17
2.2.6 What Causes Low Lattice Thermal Conductivity?	18
2.3 Temperature-Dependent Effective Potential (TDEP)	19

2.3.1	TDEP methodology	20
2.3.2	Relevance for the Articles	26
3	Results and Discussion	27
3.1	Machine Learning in Material Informatics	27
3.1.1	Introduction: Predicting DFT Volume	28
3.1.2	Features for Machine Learning in Materials Science	30
3.1.3	Machine Learning the Lattice Thermal Conductivity: Which Model is Best?	32
3.1.4	Active Sample Selection	36
3.2	Lattice Thermal Conductivity Calculations	41
3.2.1	Lattice Thermal Conductivity with Noisy Force Constants	41
3.2.2	Atomic Substitutions and Grain Boundaries Reduce Lattice Thermal Conductivity in Full-Heusler AlVFe_2	43
3.2.3	Lattice Thermal Conductivity Calculations: Comparison with Experiment	49
3.3	Limitations and Implications	51
3.3.1	Databases and Screening Alternatives	51
3.3.2	Limitations of Calculated Lattice Thermal Conductivity	53
3.3.3	Integration of Theory and Experiments	54
3.3.4	Machine Learning for Lattice Thermal Conductivity with Sublattice Substitutions	54
4	Conclusions and Outlook	57
4.1	Computational Screening for Promising Thermoelectric Compounds	57
4.2	Sublattice Substitutions and Grain Boundaries to Reduce Lattice Thermal Conductivity	59
	Bibliography	61
	Papers	78
I	Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis	79
II	Attaining Low Lattice Thermal Conductivity in Half-Heusler Sublattice Solid Solutions: Which Substitution Site Is Most Effective?	91

III	Lattice thermal conductivity of cubic compounds with iterative training over material space using Gaussian process regression	107
IV	Discarded gems: Thermoelectric performance of materials with bandgap emerging at the hybrid-functional level	125

List of Figures

1	Progress of learning throughout education, adapted from [1]. . .	xii
1.1	Left panel: Rock crystal dagger found in Spain. Photograph: Miguel Angel Blanco de la Rubia. Image is taken from Ref. [3]. Right panel: Glass window in the Augsburg Cathedral depicting King David. Photograph: Hans Bernard, CC BY-SA 3.0. Image is taken from Ref. [4]	2
2.1	Illustration of the Seebeck effect in a thermoelectric material. The arrow indicates the direction of the electric field.	6
2.2	Two numerical solutions	8
2.3	Illustration of the $Pnma$ phase for SnSe. The red atoms represent Sn and the purple atoms represent Se.	9
2.4	Phonon dispersions for Si in the diamond structure.	14
2.5	Phonon frequencies and scattering rates for full-Heusler AlVFe ₂ at 300 K. The inset illustrates three-phonon interactions.	16
2.6	The red atom is a substitute in the chain of atoms. The waves-like shapes illustrate the phonon which scatters on the substitute. . .	16
2.7	Illustration of grain boundary scattering. The dashed line indicates the grain boundary.	17
2.8	Illustration of third-order interatomic force constants. The green spheres indicate atoms, while the red and blue shaded triangles show two different three-atom interactions, each within a cutoff radius indicated by the red and blue circles. For the red triangle, atoms are displaced from their equilibrium positions. The illustration is made with inspiration from Ref. [89].	21
2.9	Number of irreducible force constants for orthorhombic ZnSb as a function of r_c . The horizontal axis indicates r_c for third-order force constants while the vertical axis shows the number of irreducible force constants.	22
2.10	Histogram showing the displacements of atoms from equilibrium positions for supercell configurations with 108 atoms at 100 K and 300 K.	24
3.1	Linear regression model for predicting the compound volume per atom. The horizontal axis shows the volume per atom calculated with DFT, and the vertical axis shows the predicted volume, V_{model} , calculated with Eq. (3.2).	29

List of Figures

3.2	Linear regression model for predicting the volume per atom. The horizontal axis shows the volume per atom calculated with DFT, and the vertical axis shows the predicted volume calculated with Eq. (3.4) (blue disks) and Eq. (3.5) (orange disks).	30
3.3	Voronoi structure illustration (turquoise) for the central atom (green).	32
3.4	Illustration of a decision tree.	33
3.5	Predictions of the LTC using RFR, SVR, and Lasso.	34
3.6	ML LTC predictions on the test set using Lasso (upper panel), RFR (lower left panel), and SVR (lower right panel). For each compound, the LTC is predicted at 200, 300, 400, 500, and 600 K.	35
3.7	Examples of distributed data. The red shape and purple shapes represent different data sets. The green crosses are sampled data points.	37
3.8	The horizontal and vertical axes show two different features (compound descriptors). The original data (blue points) are projected onto the orange line which corresponds to the first PC.	37
3.9	GPR illustration with predictions (orange line) and uncertainties (purple).	39
3.10	Illustration of compounds from active sampling using PCA and GPR STD. The left axis (red) shows the distance in PC space between the compound in the test set farthest from the training pool and its closest training pool neighbor. The right axis (blue) shows the GPR STD of the compound with the highest GPR STD in the test set.	40
3.11	Spearman correlation (left panel) and RMSE for the three compounds with the lowest TDEP LTC (right panel) for predictions made on the test set (vertical axis) vs. number of compounds sampled to the training set (horizontal axis).	40
3.12	Values of the true IFCs are shown on the horizontal axis and noisy IFCs on the vertical axis. (a) shows all IFCs, while (b) shows IFCs close to zero.	42
3.13	LTC as a function of temperature for AlVFe_2 using true IFCs (purple), with uncertainty based on calculations using IFCs with Gaussian noise and uniform noise.	43
3.14	Crystal structure of AlVFe_2	44
3.15	Phonon dispersion (left panel) and DOS (right panel) for AlVFe_2 . The projected DOS for the two equivalent Fe atoms are summed.	45
3.16	Lattice thermal conductivity at 300 K (vertical axis) for AlVFe_2 with different substitution concentrations (horizontal axis) and sites.	46
3.17	Lattice thermal conductivity for AlVFe_2 as a function of temperature. The temperature is indicated on the horizontal axis and the LTC on the vertical axis.	47

3.18	Lattice thermal conductivity at 300 K (vertical axis) with 100 nm grain boundaries for AlVFe ₂ with different substitution concentrations (horizontal axis) and sites.	48
3.19	Spectral LTC (solid lines, left vertical axis) and cumulative LTC (dashed lines, right vertical axis) as a function of phonon energy for AlVFe ₂ at 300 K.	49
3.20	Calculated $\kappa_{\ell}^{\text{TDEP}}$ (this work) versus experimental $\kappa_{\ell}^{\text{exp}}$ or κ^{exp} . For TiNiSn [66, 133–135], ZrNiSn [66], NbCoSn [136], and VFeSb [137], $\kappa_{\ell}^{\text{exp}}$ is shown on the horizontal axis, while for AgBr [138], CsI [139], MgO [140], and BP [141] κ^{exp} is shown. The data is taken at 300 K except for MgO for which it is taken at 400 K. For TiNiSn an average with standard deviation is shown where data is taken at 300 K from Refs. [66, 133] and ~ 323 K from Refs. [134, 135]	50

List of Tables

2.1	Max ZT for selected compounds based on Ref. [13].	8
-----	-------------------------------------------------------------	---

Chapter 1

Introduction

1.1 Motivation and Background

Solid-state materials have influenced the development of human societies for millennia. The importance of materials has been so substantial for certain epochs that we have labeled historical periods based on them – such as the Iron age – and in retrospect, our period will perhaps be referred to as the age of Silicon. Humans used complex crystalline materials, such as obsidian or rock crystal quartz, as tools for cutting, tanning, and showing wealth. Fig. 1.1 (left panel) displays a picture of a somewhat new (3000 BC) crystal dagger. Egyptians used iron oxide for dyeing clothes and painting their Pharaohs' tombs. Bronze – a strong combination of copper and tin – lead to a paradigm shift in farming and warfare. Transparent glass enabled light to pass through while keeping the elements at bay. As a testament to the rigidity of glass silicates, stained glass windows from 1065 have survived to this day in the Augsburg Cathedral, shown in Fig. 1.1 (right panel). The discovery and exploitation of the magnetic properties of iron alloys made land and sea traversal more efficient by enabling the construction of compasses. This invention amplified global trade and the spread of religions and facilitated knowledge transfer between scholars from all corners of the world.

What enabled the previously mentioned technologies to make a societal impact was the underlying physical properties of the materials. Understanding complex properties on the atomic scale is necessary for industrial and scientific development in the digital age. Knowledge about quantum tunneling, for example, enables transistors to work in solid-state data storage drives. The interest in novel materials with extraordinary properties is growing, and many resources have been allocated to research, e.g. through the Materials Genome Initiative [5, 6]. Research grants are often allocated to projects that aim to improve industrial processes or effectivize energy production, storage, and usage. Within energy technology, state-of-the-art materials research has enabled environmentally friendly energy harvesting. Fine-tuning of semiconductors increases solar cell efficiency making them more prominent in commercial and private use, and new battery technologies are a welcome addition to energy management and the electrification of the transport sector.

Solid-state materials do not only allow us to harvest wind energy with windmills made from steel or solar energy with silicon-based solar cells. Heat energy can be harvested with thermoelectric materials that can directly convert heat to electricity in an environmentally friendly manner. In these materials, the movement of electrons and atoms governs the efficiency of heat-to-electricity conversion. From a theoretical standpoint, the understanding of atomic-scale



Figure 1.1: Left panel: Rock crystal dagger found in Spain. Photograph: Miguel Angel Blanco de la Rubia. Image is taken from Ref. [3]. Right panel: Glass window in the Augsburg Cathedral depicting King David. Photograph: Hans Bernard, CC BY-SA 3.0. Image is taken from Ref. [4]

properties is paramount for explaining the thermoelectric effect. As of now, thermoelectric materials are typically used in niche applications, such as outer space devices, USB chargers, and stove-top fans, while large-scale applications are not as widespread. The expansion of thermoelectric technologies requires cheap, non-toxic, and efficient thermoelectric materials made from earth-abundant elements. The optimization of thermoelectric properties and the discovery of new materials are thus substantial fields of research.

The thermoelectric figure-of-merit, ZT , indicates the heat-to-electricity conversion efficacy in thermoelectric materials. ZT depends on electronic and atomic transport properties and is inversely proportional to the total thermal conductivity. The total thermal conductivity determines a crystalline material's ability to conduct heat. It consists of two parts, electronic thermal conductivity and lattice thermal conductivity. In semiconductors, the latter is typically significantly larger than the former. Top-performance thermoelectric materials, therefore, require low lattice thermal conductivity to attain a high heat-to-electricity conversion efficiency. Finding ways to reduce the lattice thermal conductivity or identifying materials with low lattice thermal conductivity can therefore be viable routes to obtain new promising thermoelectric materials.

A reliable way to find if a material has high thermoelectric efficiency is through laboratory experiments. Experiments require much expertise, equipment, and time. There are billions of different combinations of elements, and a given combination can exist in a variety of different structures. The Materials Project

database, for example, contains ~ 150000 materials [7]. Finding the best material for thermoelectric applications becomes equivalent to finding a needle in a haystack. Using computational methods – where thermoelectric properties such as the lattice thermal conductivity are estimated based on theory – can be a viable option for accelerated identification of promising thermoelectric materials [8–10]. This type of digital material search is often labeled computational screening. With screening, it is possible to analyze thousands of materials, identify promising candidates, and further condense these down to an even smaller subset based on filtering criteria. These candidates can then be proposed for experimental realization, which is the final test to verify real-world thermoelectric properties.

1.2 Thesis Introduction and Articles

The body of this thesis builds a foundation for the contents of the attached scientific articles. The thesis is partly written based on regulations and suggestions from the Faculty of Science and Technology at the Norwegian University of Life Sciences. One such suggestion is that the thesis should be readable and understandable for a new PhD candidate starting in the same field. Achieving this has been an ambition of mine while writing this thesis. The thesis is written in an effort to solidify the overarching research goals in a somewhat broad and pedagogical manner. The thesis also includes results from smaller studies that did not make it into any of the articles. The intent is that the thesis can serve as an introduction to the process of discovering thermoelectric materials and that it also promotes ideas for future research.

The main topic of this thesis is the identification of low lattice thermal conductivity materials with potential for thermoelectric applications. The lattice thermal conductivity is calculated using quantum mechanical simulations based on density functional theory. These calculations are computationally expensive needing thousands of hours on a supercomputer. To accelerate the estimation of the lattice thermal conductivity, machine learning is employed. Alloying – introducing sublattice substitutions – effectively reduces the lattice thermal conductivity and is studied separately from the machine learning-based works. The lessons learned and materials found can help guide the design of materials in the laboratory, hopefully resulting in new real-world thermoelectric applications for renewable energy conversion.

In Article I, we explore the use of machine learning for predicting the lattice thermal conductivity of half-Heusler materials. Many of these materials have favorable electronic properties – making them promising for thermoelectric applications – but they also have high lattice thermal conductivity. To gain insight into the lattice thermal conductivity as well as to obtain machine learning training data the lattice thermal conductivity of 122 half-Heusler compounds is calculated. The case study reveals that the choice of compounds used for training the machine learning model is crucial for discovering low lattice thermal conductivity materials. We provide an approach for selecting important materials to include in the training of the machine learning model, which could easily be

expanded to different material classes.

In Article II, we study the effect of alloying half-Heusler materials for reducing the lattice thermal conductivity using computational methods. Alloying means introducing an extra element into the composition, e.g. exchanging 10 % of Ge with Sn in TiNiGe, resulting in TiNiGe_{0.9}Sn_{0.1}. Alloying is often utilized to reduce the lattice thermal conductivity although the reduction varies between materials. We provide a large dataset of lattice thermal conductivity calculations with different alloying elements. We find several trends in the data, indicating both how a material should be alloyed, as well as which materials could have a drastic reduction of the lattice thermal conductivity when alloyed. These rules of thumb can guide material and alloying element selection for finding new thermoelectric materials.

Article III continues in the tracks of Article I. Here we expand the material search to a large subset of cubic materials in the Materials Project database [7]. Screening is conducted on ~ 1500 materials to separate low lattice thermal conductivity materials from the rest. Active material sampling based on Gaussian process regression identifies 30 materials with high uncertainty, which are subsequently added to the training set. We highlight 34 compounds with electronic band gaps and low lattice thermal conductivity. Na₂TlSb and Ca₃AsBr₃ – two new potential thermoelectric materials – are studied more in-depth to attain a better understanding of their lattice transport properties. The active sampling scheme and machine learning model are built in a general fashion, enabling further studies to include other material classes.

Article IV concerns screening and calculated transport properties of materials from the Materials Project database. Standard density functional theory calculations often underestimate the electronic bandgap; a key property in thermoelectric materials. In this study, we screen ~ 1000 materials and reassess those with zero band gap using a hybrid functional. Hybrid functionals have been found to often provide more accurate band structures and band gaps. Eight materials that have zero band gap with standard density functional theory have non-zero band gap with the hybrid functional. These materials are studied further and the thermoelectric figure of merit is calculated. The calculations indicate that MgSc₂Hg and Li₂CaSi have promising thermoelectric properties.

1.3 Outline

The upcoming chapters give relevant background material for the topics studied in this thesis. Chapter 2 first presents a short introduction to thermoelectricity and thermoelectric materials. Afterward is an elaboration on the theory lattice dynamics, which lays the foundation for lattice thermal conductivity calculations. Chapter 3 concerns two main topics with results and discussions: 1) Machine learning for predicting the lattice thermal conductivity, and 2) lattice thermal conductivity calculations. The chapter ends with a discussion of the limitations and implications of the work. Chapter 4 presents the conclusions, before the scientific articles of the thesis are included at the end.

Chapter 2

Materials and Methods

This chapter provides background theory and concepts relevant to the following results and discussions as well as the articles attached at the end of the thesis. First, thermoelectricity and thermoelectric materials are discussed. Second, the equations of motion for phonons are presented, and the phonon dispersions of Si are used as an example. Third, the theory of lattice thermal conductivity and phonon scattering mechanisms are outlined. Lastly, the chapter covers the temperature-dependent effective potential method which forms the basis of the lattice thermal conductivity calculations in the thesis.

2.1 Thermoelectricity: Principles and Materials

The thermoelectric effect allows for electricity to be harnessed from heat as well as the reverse process; using electricity for heating and cooling. In thermoelectrics, the heat-electricity conversion occurs without any moving parts. Thermoelectric generators can thus be essentially maintenance free and noiseless. The thermoelectric effect is used for cooling in car seats, wine refrigerators, and scientific equipment. It is also used for heat-to-electricity conversion in wood stove fans, thermometers, and space rovers. Applications of thermoelectrics can involve wearable devices that harness body heat to charge cell phones or medical and monitoring sensors connected to the "Internet of Things" [11]. There are also good prospects to employ thermoelectrics in larger scale industries that generate vast amounts of heat, e.g. in metal furnaces, where waste heat is often used for heating water or simply let out into the surroundings.

One aspect limiting the use of thermoelectric materials in large-scale waste heat regeneration units is their relatively low efficiency. Thermoelectric materials are typically therefore still mostly used in niche products. Nonetheless, efforts have been made to improve thermoelectric materials, and record-breaking thermoelectric efficiencies have been reported over the last decades [12–14]. This research shows that there are good prospects to improve thermoelectric efficiency, both in terms of optimizing known materials as well as discovering new ones. High-performance materials often contain toxic elements such as Pb or Te, or rare earth elements [15]. It is a focus in the research community to make efficient thermoelectric materials that contain non-toxic, Earth-abundant elements in accordance with rules based on environmental laws set in place to protect the environment, product users, and workers [16]. This section covers basic concepts of thermoelectricity, such as the Seebeck effect and the Peltier effect. It ends with a discussion on high-performance thermoelectric materials from the literature.

2.1.1 Seebeck and Peltier Effect

Certain materials attain a voltage difference when there is a temperature difference across the material. This property is known as the Seebeck effect, named after the scientist who discovered it circa 200 years ago, Thomas Johann Seebeck. The discovery led to the definition of the Seebeck coefficient, α . This coefficient represents a material's ability to sustain a voltage, V , when there is a temperature difference, ΔT , across the material,

$$\alpha = \frac{V}{\Delta T}. \quad (2.1)$$

When a material with non-zero α has a temperature difference across it, mobile charge carriers diffuse between the hot side to the cold side. This in turn results in an electric field within the material because of a higher charge density on one side than the other, as illustrated in Fig. 2.1.

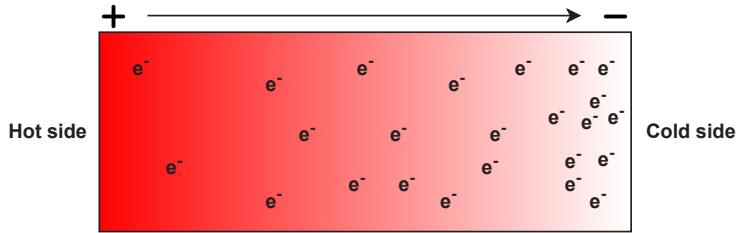


Figure 2.1: Illustration of the Seebeck effect in a thermoelectric material. The arrow indicates the direction of the electric field.

Jean Charles Athanase Peltier discovered that a current passing through a material could change the temperature locally, initiating a temperature difference on each side. The effect was dubbed the Peltier effect. This effect gives name to the Peltier cooler, a device that can use electricity for local cooling and heating. As an analogue to the Seebeck coefficient, the Peltier coefficient is defined as

$$\Pi = \frac{Q}{I}, \quad (2.2)$$

where Q is the heat generated per unit of time and I is the current passing through the material. Materials exhibiting the Seebeck effect – and the reverse Peltier effect – are called thermoelectric materials.

2.1.2 Thermoelectric Figure of Merit

A thermoelectric generator harvests electric energy from a temperature gradient in a thermoelectric material. The maximum heat-to-electricity conversion efficiency is given by [17]

$$\eta_{\max} = \frac{T_h - T_c}{T_h} \frac{\sqrt{1 + ZT_{\text{avg}}} - 1}{1 + ZT_{\text{avg}} + T_c/T_h}, \quad (2.3)$$

where T_h is the temperature on the hot side, T_c is the temperature on the cold side, $T_{\text{avg}} = (T_h + T_c)/2$, and Z is the thermoelectric figure of merit. Eq. 2.3 is limited by the assumption that Z is temperature independent, and is therefore most accurate when the temperature difference is small.

The dimensionless figure of merit (ZT) of a thermoelectric material conventionally determines the thermoelectric energy conversion efficiency. This quantity enables comparison of the conversion efficiency of different thermoelectric materials, where Z can be taken as temperature dependent. The dimensionless figure of merit is defined as

$$ZT = \frac{\alpha^2 \sigma T}{\kappa} = \frac{\mathcal{P}T}{\kappa}. \quad (2.4)$$

Here, α is the Seebeck coefficient, σ is the electrical conductivity, $\mathcal{P} = \sigma S^2$ is the power factor, T is the absolute temperature, and κ is the thermal conductivity. The thermal conductivity is divided into two parts, $\kappa = \kappa_\ell + \kappa_e$, where κ_ℓ is the lattice thermal conductivity (LTC) and κ_e is the electronic thermal conductivity. Obtaining a high ZT requires that the thermoelectric material has high α and σ , and low κ . A high α corresponds to a large voltage difference as seen from Eq. 2.1. A low κ means that heat energy is transferred slowly and that a high temperature gradient can be maintained. ZT is made up of interdependent parameters; optimizing one can lead to changes in the others. An example of this is seen with the Wiedemann-Franz law, which states that the electrical conductivity and electronic thermal conductivity are proportional, $\kappa_e = \sigma LT$, where L is the Lorentz number. An optimal charge carrier density thus maximizes the electronic part of ZT . The LTC is decoupled from the Wiedemann-Franz law – a low LTC is thus required for attaining high ZT .

2.1.3 Thermoelectric Materials

As discussed above, there are different material property requirements for attaining a high ZT . Typically, the best-performing thermoelectric materials are doped semiconductors. Doped semiconductors can maintain a high α and σ at the same time [12]. The value of α is tied to the electronic structure and bandgap, E_g . In metals, which have $E_g = 0$, α is low. This typically makes metals poor thermoelectric materials, although they can have large σ . Semi-metals with properties in-between that of metals and semiconductors can attain suitable charge carrier densities and have shown promising thermoelectric performance [18–20].

Fig. 2.2 shows ZT of high-performance thermoelectric materials taken from Ref. [12]. The shape of the ZT curves dictates the optimal temperature range for a material. The maximum ZT ranges from ~ 0.5 (p-type SiGe) to ~ 1.2 (p-type TAGS). For low temperatures, Bi_2Te_3 (n-type) and Sb_2Te_3 (p-type) are most efficient, and for high temperatures, SiGe (n-type) and $\text{Yb}_{14}\text{MnSb}_{11}$ (p-type) come out on top. It is worth noting that it is not solely ZT that decides the optimal material to put in a device from an engineering standpoint. Other factors such as longevity and toxicity also impact what material is most suited for a

2. Materials and Methods

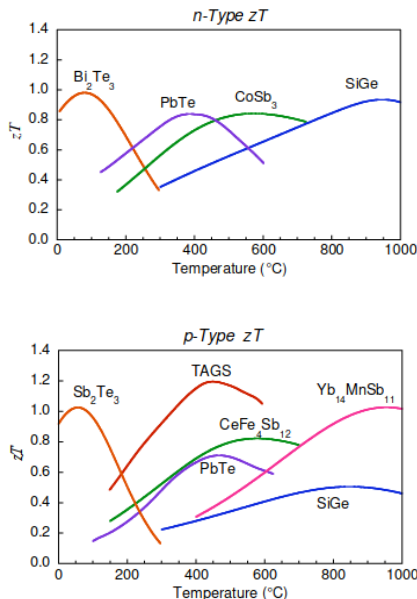


Figure 2.2: Thermoelectric figure of merit for high-performance thermoelectric materials. The top panel shows n-type materials and the bottom panel shows p-type materials. The plots are taken from Ref. [12]

Table 2.1: Max ZT for selected compounds based on Ref. [13].

	SnSe	$\text{Ta}_{0.74}\text{V}_{0.1}\text{Ti}_{0.16}\text{FeSb}$	$\text{La}_2\text{Ti}_2\text{O}_7$	$\text{Ge}_{0.9}\text{Sb}_{0.1}\text{Te}_{1.03}$
ZT	3.1 [21]	1.5 [22]	2.6 [23]	1.9 [24]

given application. For example, lead could potentially be used in thermoelectric materials for space probes but could be less desirable in wearable thermoelectric devices.

Efforts have been made to achieve high ZT in different material types and classes; SnSe [25, 21, 26], Cu_2Se [27, 28], ZnO [29, 30], half-Heuslers [31–33], full-Heuslers [34, 35], and perovskites [36, 37], to name a few. A summary of high-performance materials based on Ref. [13] is shown in Table. 2.1. SnSe has recently shown its promise as a highly efficient thermoelectric material [25]. For a polycrystalline sample $ZT = 3.1$ at 783 K was reported from experiments [21]. The high ZT was a result of favorable electronic properties and an ultra-low LTC of 0.07 W/Km. An illustration of the room temperature $Pnma$ phase for SnSe is shown in Fig. 2.3. There is a large variety in the structural types of thermoelectric materials, e.g. layered and cubic, monocrystalline and polycrystalline, and non-oxide and oxide materials. The large variety of material types combined with

the vast number of possible element combinations makes up a large materials space with potential novel thermoelectric materials.

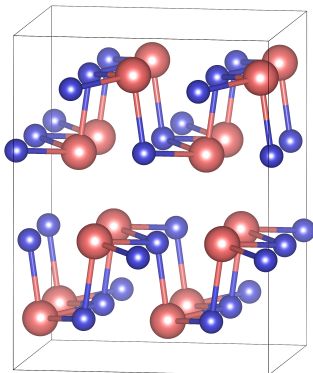


Figure 2.3: Illustration of the $Pnma$ phase for SnSe. The red atoms represent Sn and the purple atoms represent Se.

2.1.4 Relevance for the Articles

Article I:

This article consists of a case study exploring how to reliably identify low LTC compounds using machine learning. The study includes an analysis of the half-Heusler class of compounds. Among the half-Heuslers, there are several promising thermoelectric candidates [31, 38]. Screening for low LTC can be a viable option to discover new efficient thermoelectric materials. The screening schemes in the article are extendable to other classes of compounds. The machine learning predicted LTC can be obtained without extensive calculations and can potentially be employed for large compound sets.

Article II:

In this article, we also study the half-Heusler compounds. Here partial isovalent sublattice substitutions and grain boundaries are introduced in calculations to reduce the LTC. The half-Heuslers have been shown to have favorable electronic transport properties, while having relatively high intrinsic LTC, limiting their attainable ZT [39]. The calculations reveal that several of the half-Heuslers with high intrinsic LTC can have a drastic reduction of the LTC with the introduction of additional phonon scattering mechanisms. The findings in the article could be adopted to other compound classes which in turn could yield promising thermoelectric compounds with low LTC and high ZT .

Article III:

In this work we train a machine learning model to predict the LTC of cubic compounds from the Materials Project database [7]. In the screening study we include metals, semiconductors, and insulators. We find known TE compounds, such as Sr- and Ba-based full-Heuslers. We also find new low LTC semiconductors, where Na_2TlSb and Ca_3AsBr_3 are highlighted. These compounds can be explored

2. Materials and Methods

further in continuing studies including the calculation of the electronic properties for obtaining ZT .

Article IV: This work contains a screening study of compounds with four atoms in the primitive cell, where the electronic transport properties are calculated for compounds with band gaps emerging with a hybrid functional. Calculations show peak $ZT \sim 1$ for the full-Heuslers MgSc_2Hg and Li_2CaSi . These compounds also display relatively low $\text{LTC} \leq 5 \text{ W/Km}$ at 300 K. While the toxicity of Hg unfortunately reduces the applicability of MgSc_2Hg in thermoelectric devices, Li_2CaSi contains non-toxic elements.

2.2 Lattice Dynamics

The topic of lattice dynamics involves understanding atomic vibrations in a crystalline solid. The collective movement of atoms, called phonons, participate in a number of material phenomena in solids [40], such as heat flow, thermal expansion, phase transitions, and propagation of sound. Phonon behavior is affected by several material properties — temperature, interatomic distances, atomic bonding, and defects to name a few. This section presents a basic formalism of lattice dynamics and the harmonic approximation of the lattice potential energy with notation based on the book *Physics of Phonons* [40]. The phonon dispersions of Si are presented to link theory and practice. The section also outlines the theory of lattice thermal conductivity and phonon scattering mechanisms. Lastly, potential routes for achieving low lattice thermal conductivity are discussed.

2.2.1 Lattice Potential Energy and Phonons

To obtain an expression for the lattice potential energy it is assumed that the potential energy of the crystal lattice depends on atomic displacements from equilibrium. A distortion of atoms away from the equilibrium positions should increase the potential energy. In the following $\mathbf{u}(bl)$ is the displacement of atom b in unit cell l , and summations are over cells and atoms. The potential energy, U , can be expressed as a Taylor expansion,

$$\begin{aligned}
 U = U_0 &+ \sum_{lb\alpha} \left. \frac{\partial U}{\partial u_\alpha(bl)} \right|_0 u_\alpha(bl) + \frac{1}{2!} \sum_{lb,l'b'} \sum_{\alpha\beta} \left. \frac{\partial^2 U}{\partial u_\alpha(bl) \partial u_\beta(b'l')} \right|_0 u_\alpha(bl) u_\beta(b'l') \\
 &+ \frac{1}{3!} \sum_{lb,l'b',l''b''} \sum_{\alpha\beta\gamma} \left. \frac{\partial^3 U}{\partial u_\alpha(bl) \partial u_\beta(b'l') \partial u_\gamma(b''l'')} \right|_0 u_\alpha(bl) u_\beta(b'l') u_\gamma(b''l'') + \dots,
 \end{aligned} \tag{2.5}$$

where U_0 is the equilibrium potential energy which can be set to zero for the dynamical problem, and α , β , and γ are Cartesian coordinates. The second term on the right-hand side corresponds to a force that is zero when evaluated at equilibrium. Within the harmonic approximation, the Taylor series is expanded to the second-order derivative,

$$U_{\text{harm}} \equiv \frac{1}{2} \sum_{bl,b'l'} \sum_{\alpha\beta} \left. \frac{\partial^2 U}{\partial u_\alpha(bl) \partial u_\beta(b'l')} \right|_0 u_\alpha(bl) u_\beta(b'l'). \tag{2.6}$$

The next step is to introduce a term representing the force on atom bl in direction α when atom $b'l'$ is displaced in direction β using the second derivative of the potential energy evaluated at equilibrium,

$$\Phi_{\alpha\beta}(bl, b'l') = \left. \frac{\partial^2 U}{\partial u_\alpha(bl) \partial u_\beta(b'l')} \right|_0. \tag{2.7}$$

The equations of motion then become

2. Materials and Methods

$$m_b \ddot{u}_\alpha(bl) = - \sum_{b'l'\beta} \Phi_{\alpha\beta}(bl, b'l') u_\beta(b'l'), \quad (2.8)$$

where m_b is the mass of atom b . The force matrix, Φ , has several symmetry relations among which is the lattice translational symmetry. This allows translating the unit cell by l ,

$$\Phi_{\alpha\beta}(bl, b'l') = \Phi_{\alpha\beta}(b0, b'(l' - l)). \quad (2.9)$$

Combining Eq. 2.8 and Eq. 2.9 the equations of motion become

$$m_b \ddot{u}_\alpha(bl) = - \sum_{b'l'\beta} \Phi_{\alpha\beta}(b0, b'l') u_\beta(b'l'). \quad (2.10)$$

As a solution to Eq. 2.10 a plane wave ansatz is used,

$$u_\alpha(bl) = \frac{1}{\sqrt{m_b}} \sum_{\mathbf{q}} U_\alpha(\mathbf{q}, b) \exp[i(\mathbf{q} \cdot \mathbf{x}(l) - \omega t)], \quad (2.11)$$

where $\mathbf{x}(l)$ is the equilibrium position of unit cell l , \mathbf{q} is a wave vector, $U_\alpha(\mathbf{q}, b)$ is the amplitude, and ω is the frequency. Combining Eq. 2.10 and Eq. 2.11 we arrive at

$$\omega^2 U_\alpha(\mathbf{q}b) = \sum_{b'\beta} D_{\alpha\beta}(bb'|\mathbf{q}) U_\beta(\mathbf{q}, b), \quad (2.12)$$

where the representation of the dynamical matrix is

$$D_{\alpha\beta}(bb'|\mathbf{q}) = \frac{1}{\sqrt{m_b m_{b'}}} \sum_{l'} \Phi_{\alpha\beta}(b0, b'l') \exp(i\mathbf{q} \cdot \mathbf{x}(l')). \quad (2.13)$$

The solutions of Eq. 2.12 can be obtained when the following determinant is zero,

$$|D_{\alpha\beta}(bb'|\mathbf{q}) - \omega^2 \delta_{\alpha\beta} \delta_{bb'}| = 0, \quad (2.14)$$

where δ_{ij} is the Kronecker delta.

With this equation, solutions on the form $\omega = \omega(\mathbf{q}s)$ are sought, corresponding to vibrational frequencies at wave vector \mathbf{q} and branch s . The number of phonon branches depends on the number of atoms in the unit cell, p , such that $s = 1, 2, 3 \dots, 3p$. For the $3p$ eigenvalues $\omega^2(\mathbf{q}s)$ at \mathbf{q} there is an eigenvalue equation from rewriting Eq. 2.12,

$$\omega^2(\mathbf{q}s) \epsilon_\alpha(b, \mathbf{q}s) = \sum_{b'\beta} D_{\alpha\beta}(bb'|\mathbf{q}) \epsilon_\beta(b', \mathbf{q}s), \quad (2.15)$$

with eigenvectors $\epsilon(b, \mathbf{q}s)$. The collection of $\omega(\mathbf{q}s)$ is called the phonon dispersions. The frequencies have periodicity in the same manner as the reciprocal lattice, and solutions can be limited to \mathbf{q} within the first Brillouin zone. Three

branches, called acoustic, are such that $\omega(\mathbf{q}) \rightarrow 0$ for $\mathbf{q} \rightarrow 0$, and $3p - 3$ branches, called optical, are such that $\omega(\mathbf{q}) \neq 0$ for $\mathbf{q} \rightarrow 0$. If a solution to Eq. 2.14 results in an eigenvalue $\omega^2 < 0$, it is a sign that the crystal is dynamically unstable. For $\text{Im}(\omega) \neq 0$ the displacements from equilibrium will grow exponentially, seen from the time-dependent part of the displacement oscillations, $\exp(-i\omega t)$, of Eq. 2.11. In this case, the energy decreases when atoms are moved away from equilibrium.

2.2.2 Phonon Dispersions of Si

Fig. 2.4 shows the phonon dispersion relations for Si in a diamond lattice. Γ , X , W , K , L , and U are high symmetry points in the Brillouin zone. Γ corresponds to $\mathbf{q} = 0$ and long wavelength phonons. The primitive cell contains two Si atoms, giving rise to three acoustic and three optical branches. For the acoustic branches when $\mathbf{q} \rightarrow 0$ the two atoms in the primitive cell vibrate in the same direction, while in the optical branches, the atoms vibrate against each other. Acoustic and optical branches can be transversal or longitudinal. In transversal motion, the atoms move perpendicular to the direction of the wave, \mathbf{q} , while for longitudinal motion, the atoms move parallel to \mathbf{q} . Close to Γ there is two transversal acoustic (TA) branches and one longitudinal acoustic (LA) branch.

The velocity of the acoustic phonons in the long-wavelength limit, $\mathbf{q} \rightarrow 0$, is called the speed of sound, connecting microscopic and macroscopic properties. The phonon group velocity can be calculated as $v_{\mathbf{q}s} = d\omega_{\mathbf{q}s}/d\mathbf{q}$. $v_{\mathbf{q}s}$ for the LA branch is typically higher than for the TA branch close to $\mathbf{q} = 0$. The calculated $v_{\mathbf{q}s}$ for the three acoustic branches for Si close to the Γ -point in the direction of X are $v_{\mathbf{q}s}^{\text{TA1}} = 5192$ m/s, $v_{\mathbf{q}s}^{\text{TA2}} = 5192$ m/s, and $v_{\mathbf{q}s}^{\text{LA}} = 7441$ m/s. These values can be compared to previous experimental estimates, calculated based on the elastic tensor [41, 42], $v_{\mathbf{q}s}^{\text{TA}} = 5843$ m/s and $v_{\mathbf{q}s}^{\text{LA}} = 8433$ m/s. There is a difference between the velocities, but the ratio between the LA and TA velocities is similar at ~ 1.4 . The velocities for the two TA modes between Γ and X are identical as the dispersions are overlapping. Such branches are often labeled as degenerate. Close to the Γ -point, the optical branches are flatter than the acoustic branches, and the phonon group velocity of the optical phonons is lower than the group velocity of the acoustic phonons in this region.

2.2.3 Lattice Thermal Conductivity

The lattice thermal conductivity (LTC) expresses the rate of heat transport per unit temperature due to the phonons. The heat transport depends on the phonon frequency, ω_λ , the phonon group velocity, v_λ , and the non-equilibrium phonon distribution, $n_{\lambda\alpha}$, where $\lambda = (\mathbf{q}s)$, and α is a Cartesian direction. With a temperature gradient, ∇T_α , and solid volume, V , the heat conducting phonon current can be expressed as

$$J_\alpha = \frac{1}{V} \sum_\lambda \hbar \omega_\lambda v_{\lambda\alpha} n_{\lambda\alpha}. \quad (2.16)$$

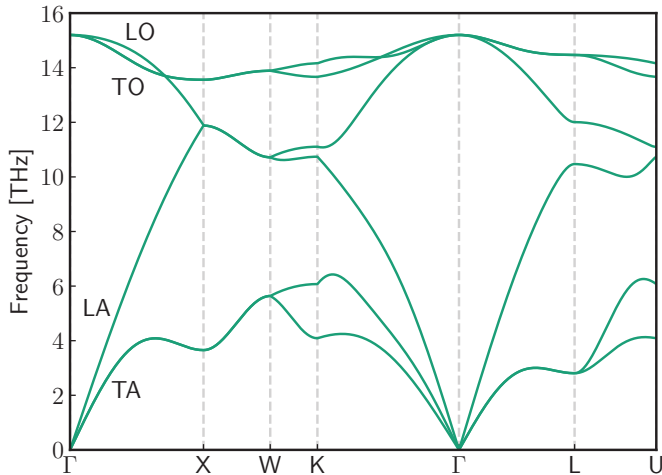


Figure 2.4: Phonon dispersions for Si in the diamond structure.

The non-equilibrium phonon distribution can be approximated as a deviation from the equilibrium phonon distribution, \tilde{n}_λ . With a small temperature gradient the phonon distribution becomes [40]

$$n_{\lambda\alpha} \approx \tilde{n}_\lambda - v_{\lambda\alpha} \tau_{\lambda\alpha} \frac{d\tilde{n}_\lambda}{dT} \frac{dT}{d\alpha}, \quad (2.17)$$

where $\tau_{\lambda\alpha}$ is the phonon lifetime. The equilibrium distribution does not contribute to the heat transport, such that when inserting Eq. (2.17) into Eq. (2.16) the phonon current becomes

$$J_\alpha = -\frac{1}{V} \sum_\lambda \hbar\omega_\lambda \frac{d\tilde{n}_\lambda}{dT} v_{\lambda\alpha} v_{\lambda\alpha} \frac{dT}{d\alpha}. \quad (2.18)$$

Here, the phonon heat capacity is $c_\lambda = \hbar\omega_\lambda d\tilde{n}_\lambda/dT$. Combining Eq. (2.18) with Fourier's law, $J = -\kappa \nabla T$, we arrive at an expression for the LTC,

$$\kappa_{\alpha\beta} = \frac{1}{V} \sum_\lambda c_\lambda v_{\lambda\alpha} v_{\lambda\beta} \tau_{\beta\lambda}. \quad (2.19)$$

This tensor contains nine elements,

$$\kappa = \begin{pmatrix} \kappa_{xx} & \kappa_{xy} & \kappa_{xz} \\ \kappa_{yx} & \kappa_{yy} & \kappa_{yz} \\ \kappa_{zx} & \kappa_{zy} & \kappa_{zz} \end{pmatrix}, \quad (2.20)$$

The off-diagonal elements correspond to heat conduction that is non-parallel to the temperature gradient. The inverse of the phonon lifetime, $1/\tau_{\beta\lambda}$, is the

phonon scattering rate. Phonon scattering mechanisms, such as three-phonon and defect scattering, are detailed in the next section.

2.2.4 Three-Phonon, Mass-Disorder, and Grain-Boundary Scattering

If phonons are taken as fully non-interacting, phonon scattering would be zero and the LTC infinite. Three-phonon scattering events cause finite LTC in a perfect lattice. Fig. 2.5 shows an example of calculated three-phonon scattering rates as a function of frequency. The number of points corresponds to the number of q -points used to discretize the Brillouin zone in Eq. (2.19). The scattering rate is given by

$$\frac{1}{\tau_\lambda} = \frac{\hbar\pi}{8} \sum_{\lambda'\lambda''} |\Phi_{\lambda\lambda'\lambda''}|^2 [(n_{\lambda'} + n_{\lambda''} + 1)\delta(\omega_\lambda - \omega_{\lambda'} - \omega_{\lambda''}) + 2(n_{\lambda'} - n_{\lambda''})\delta(\omega_\lambda - \omega_{\lambda'} + \omega_{\lambda''})]. \quad (2.21)$$

Here, the δ -functions ensure energy and momentum conservation in three-phonon interactions, $\mathbf{q} \pm \mathbf{q}' + \mathbf{q}'' = \mathbf{G}$ and $\omega \pm \omega' = \omega''$. These interactions represent two possible cases, one where two phonons combine to a third and one where one phonon decays into two, as illustrated in the inset of Fig. 2.5. In the following, the atom and cell indices, b and l , are contracted to a single index, i , denoting an atom. The three-phonon matrix element, $\Phi_{\lambda\lambda'\lambda''}$, is given as

$$\Phi_{\lambda\lambda'\lambda''} = \sum_{ijk} \sum_{\alpha\beta\gamma} \frac{\epsilon_{\alpha i}^\lambda \epsilon_{\beta j}^{\lambda'} \epsilon_{\gamma k}^{\lambda''}}{\sqrt{m_i m_j m_k} \sqrt{\omega_\lambda \omega_{\lambda'} \omega_{\lambda''}}} \times \Phi_{ijk}^{\alpha\beta\gamma} e^{i(\mathbf{q}\mathbf{r}_i + \mathbf{q}'\mathbf{r}_j + \mathbf{q}''\mathbf{r}_k)}. \quad (2.22)$$

Here, ijk are atom indices, ϵ^λ is the eigenvector of mode λ , and $\Phi_{ijk}^{\alpha\beta\gamma}$ are the third-order force constants. Details on the third-order force constants and how to obtain them are elaborated in Sec. 2.3.

There are additional physical mechanisms that cause phonon scattering, such as sublattice substitutions. A substituent atom has a different mass and bonding to its neighbors. An illustration of an atomic substitution is shown in Fig. 2.6. One way to include the effect of substitutions in the phonon scattering rate is through mass-disorder (md) scattering. In this model, the mass of the atom is changed to an effective average mass. E.g. when substituting Pb by Sn in $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ the effective mass of Pb becomes $(1-x)\%$ the mass of Pb plus $x\%$ the mass of Sn. The difference in mass is reflected in the mass-variance parameter,

$$g_i = \sum_p c_i^p \left(\frac{m_i^p - \bar{m}_i}{\bar{m}_i} \right)^2. \quad (2.23)$$

2. Materials and Methods

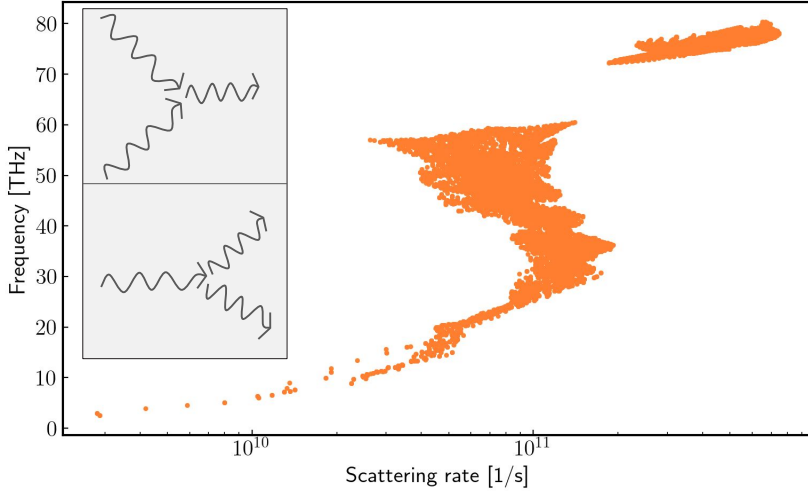


Figure 2.5: Phonon frequencies and scattering rates for full-Heusler AlVFe₂ at 300 K. The inset illustrates three-phonon interactions.

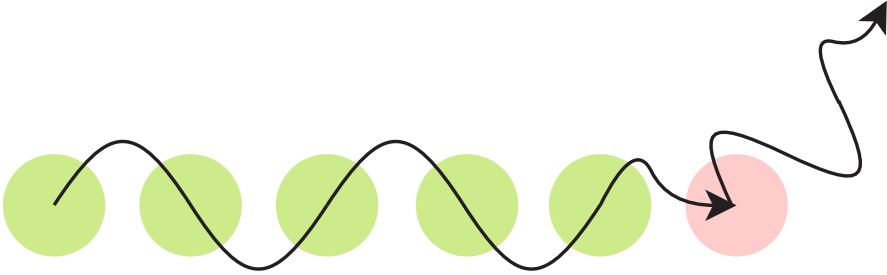


Figure 2.6: The red atom is a substitute in the chain of atoms. The waves-like shapes illustrate the phonon which scatters on the substitute.

Here, atom i is substituted by atom p at concentration c_i^p , m_i^p is the mass of atom p , and $\bar{m}_i = \sum_p c_i^p m_i^p$ is the average atomic mass. The md-scattering rate is given as [43, 44]

$$\frac{1}{\tau_\lambda^{\text{md}}} = \frac{\pi}{2} \sum_{\lambda'} \omega_\lambda \omega_{\lambda'} \sum_i g_i |\epsilon_\lambda^{i\dagger} \cdot \epsilon_{\lambda'}^i|^2 \delta(\omega_\lambda - \omega_{\lambda'}), \quad (2.24)$$

where the sum over i goes over all atoms in the cell. When no substitutions are included, $1/\tau_\lambda^{\text{md}}$ only contains the effect of atomic isotopes. The md-scattering increases for higher frequencies, and typically impacts high frequency acoustic phonons [45]. An example of how md-scattering reduces the LTC is shown in Sec. 3.2.2.

Grain boundaries affect phonon transport through grain-boundary (gb) scattering mechanisms. Grain boundaries in polycrystalline compounds cause phonons of different wavelengths to scatter, as illustrated in Fig. 2.7. The gb shown is between two differently oriented lattices indicated by red and purple discs.

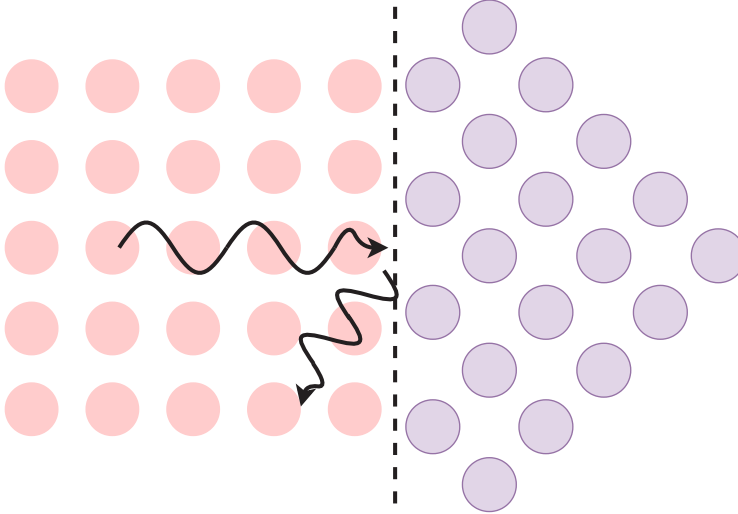


Figure 2.7: Illustration of grain boundary scattering. The dashed line indicates the grain boundary.

Calculating gb-scattering on an atomic level is challenging, as grains with different orientations and sizes can form in a variety of different ways. The size of the grains also poses a modeling challenge from a first-principles point of view. An approximation of the gb-scattering can be obtained with a computationally efficient diffusive scattering model. In this model phonons with mean-free paths comparable to the grain size are absorbed and re-emitted in the scattering process. The gb-scattering rate is given as [40]

$$\frac{1}{\tau_{\lambda}^{\text{gb}}} = \frac{v_{\lambda}}{d}, \quad (2.25)$$

where v_{λ} is the phonon group velocity and d is the grain size. The grain size can be tuned for a specific compound according to the average grain size found in experiment, or set to a smaller value as an upper bound on the scattering rate due to gbs. The gb-scattering typically suppresses low-frequency acoustic phonons. Sec. 3.2.2 shows an example of how gb-scattering lowers the LTC.

2.2.5 Lattice Thermal Conductivity Models

Calculating the LTC with the three-phonon scattering rate is computationally expensive when based on density functional theory (DFT) calculations. There

2. Materials and Methods

are alternative models that can be used to calculate the LTC which are less expensive. One model is the Slack model for the LTC [46],

$$\kappa_{\ell, \text{Slack}} = A \frac{M_a \mu n^{1/3} \theta_D^3}{\gamma^2 T}. \quad (2.26)$$

Here, $A = 2.34 \times 10^{-8} / (1 - 0.514/\gamma + 0.28/\gamma^2)$, M_a is the average atomic mass, μ^3 is the volume per atom, n is the number of atoms in the primitive cell, θ_D is the Debye temperature, and γ is the Grüneisen parameter. Another model which can be used to estimate the minimum LTC is the Cahill model [47],

$$\kappa_{\ell, \text{Cahill}} = \left(\frac{\pi}{6}\right) k_B n^{2/3} \sum_i v_i \left(\frac{T}{\theta_{D,i}}\right)^2 \int_0^{\theta_{D,i}/T} \frac{x^3 e^x}{(e^x - 1)^2} dx, \quad (2.27)$$

where the sum is taken over two transverse and one longitudinal branch, with v_i and θ_i being the group velocity and Debye temperature for branch i . The Debye temperature and Grüneisen parameter can be obtained from the phonon dispersion relations. The Debye temperature can be expressed as being the temperature associated with the highest frequency of a normal vibration mode, $\omega_{i, \text{max}}$, by the relation $\theta_{D,i} k_B = \hbar \omega_{i, \text{max}}$. The Grüneisen parameter can be obtained from dispersions at different volumes, $\gamma_i = (-V/\omega_i)(\partial\omega_i/\partial V)$, and characterizes the relationship between changes in phonon frequency and volume. The Debye temperature and Grüneisen parameter can also be estimated based on the elastic tensor from DFT [48, 49]. How to calculate the Debye temperature using the elastic tensor is shown in Sec. 2.3.

2.2.6 What Causes Low Lattice Thermal Conductivity?

When analyzing the LTC it can be useful to note the units of the LTC, typically given as $\text{W/mK} \sim \text{J/smK}$. The LTC is the energy transferred per second, per meter, per Kelvin. From Eq. (2.19) there are three central properties that affect the LTC: the heat capacity, the phonon group velocity, and the relaxation time. These properties are not independent of each other, such that altering one could affect the others. If the heat capacity is low, the phonons carry less heat energy. If the group velocity is low, the phonons transfer the heat at a slower rate. If the relaxation time is low, phonons scatter more frequently, limiting their ability to transfer heat.

Three-phonon scattering events effectively reduce the phonon lifetimes. Compounds with high three-phonon scattering rates are often labeled as anharmonic compounds. PbTe is a high-performance thermoelectric material, where its low LTC has been attributed to its high anharmonic (three-phonon) scattering rate [50]. The same is the case for the well-known thermoelectric GeTe [51]. The heightened phonon scattering in this compound has been linked to its proximity to a structural phase transition. Pal et al. found a class of quaternary chalcogenides with high scattering rates [52], showing that a large variety of compounds have this attribute.

Rattling is a phenomenon that has been linked to low LTC. A rattling atom is loosely defined as a weakly bound atom with long atomic displacements [53]. The rattling atoms can be thought of as filler atoms in cage-like voids in the structure, e.g. as seen in clathrates, giving rise to flat phonon branches [54]. Pal et al. found a class of rattling quaternary chalcogenides [52]. They observed strongly localized phonons where few atoms in the cell contributed to the vibrations. Rattling atoms can facilitate low phonon velocities or high scattering and have also been found in low LTC skutterudites [55] and half-Heusler compounds [56].

Increasing the total scattering rate by introducing substitutions can reduce the LTC. Substituted atoms increase sublattice disorder and act as local defects as illustrated in Fig. 2.6. Sublattice disorder can be introduced by partially substituting an element with another element of the same periodic group maintaining the same valence. These substitutions yield high mass variance and can be used as an efficient strategy where different substitution concentrations are investigated to optimally reduce the LTC [45, 57–60].

Doping is typically done to enhance the electronic properties of thermoelectric compounds. Although done at low concentrations, doping slightly increases sublattice disorder. The dopant atoms have a different atomic sizes, which could introduce strain in the lattice. Doping can lead to lower LTC [61], although in other cases doping increases both the LTC and the total thermal conductivity [62].

Enhancing phonon scattering with grain boundaries is also a viable route for reducing LTC. Grain boundaries emerge in polycrystalline compounds and effectively scatter phonons as illustrated in Fig. 2.7. Fine-tuning of the grain sizes has been studied both theoretically and experimentally, such as for ZnO [29, 63], half-Heuslers [45, 64–66], and perovskites [67, 68]. The sizes of the grain boundaries impact phonon transport – smaller grain boundaries facilitate more scattering. Grain boundary sizes observed in experiments vary between compounds and manufacturing methods. For half-Heuslers they have been observed to be in the range of 60 to 400 nm [64, 66].

Several other scattering mechanisms impact the LTC, such as vacancies [69–71], electron-phonon interactions [72, 73], dislocations [74], precipitates [75], pores [76], and twin boundaries [77, 78]. The relationship between these scattering mechanisms and lattice thermal transport is outside the scope of this thesis. Lastly, it should be noted that for thermoelectric compounds, the goal is to attain high ZT , not just low LTC. If the introduction of additional phonon scattering mechanisms negatively affects electronic transport properties it could result in a net lowering of ZT .

2.3 Temperature-Dependent Effective Potential (TDEP)

To calculate the LTC as presented in the previous section, Eq. (2.19), third-order force constants are needed. These force constants correspond to a third-order expansion of the lattice potential energy Taylor series. Several softwares can be used in the process of obtaining force constants from ab initio calculations,

2. Materials and Methods

such as Phono3py [79, 80], ShengBTE [81], and hiPhive [82]. One method for doing finite temperature lattice dynamics is the temperature-dependent effective potential (TDEP) method as developed by Hellman et. al [83–85]. The TDEP package contains programs for extracting second-, third-, or higher-order effective force constants. It can also be used for calculating properties such as phonon dispersions, free energies, and the LTC including different phonon scattering mechanisms.

Throughout this work, TDEP uses atomic positions and forces calculated with DFT as input. But in principle, any method that provides these properties, such as force fields or machine learning models, can be used with TDEP. Most of the theory written out in the section is based on Refs. [83–88], as well as the PhD thesis by Nina Shulumba [89], which gives an in-depth elaboration on lattice dynamics and the TDEP method. It should also be mentioned that the examples in the online TDEP manual [90] have also been of much use throughout the work of this thesis.

2.3.1 TDEP methodology

The TDEP method mainly works as a tool to obtain a model Hamiltonian for doing lattice dynamics. When expanding the Hamiltonian to the third order the following expression is obtained,

$$\hat{H} = U_0 + \sum_{i\alpha} \frac{p_{i\alpha}^2}{2m_i} + \frac{1}{2!} \sum_{i,j} \sum_{\alpha,\beta} \Phi_{ij}^{\alpha\beta} u_i^\alpha u_j^\beta + \frac{1}{3!} \sum_{i,j,k} \sum_{\alpha,\beta,\gamma} \Phi_{ijk}^{\alpha\beta\gamma} u_i^\alpha u_j^\beta u_k^\gamma. \quad (2.28)$$

Here, the first two terms on the right-hand side are the potential energy of the static lattice and the kinetic energy of the atoms. The displacement from equilibrium of atom i in direction α is indicated by u_i^α , and $\Phi_{ij}^{\alpha\beta}$ and $\Phi_{ijk}^{\alpha\beta\gamma}$ are the second- and third-order force constants representing two- and three-body atomic interactions. These force constants are often referred to as harmonic and anharmonic force constants. For a two-body interaction, the force constant matrix contains 3^2 variables, as shown in the following expression,

$$\Phi_{ij} = \begin{pmatrix} \Phi_{ij}^{\alpha\alpha} & \Phi_{ij}^{\alpha\beta} & \Phi_{ij}^{\alpha\gamma} \\ \Phi_{ij}^{\beta\alpha} & \Phi_{ij}^{\beta\beta} & \Phi_{ij}^{\beta\gamma} \\ \Phi_{ij}^{\gamma\alpha} & \Phi_{ij}^{\gamma\beta} & \Phi_{ij}^{\gamma\gamma} \end{pmatrix}, \quad (2.29)$$

while for three-body interaction the force constant matrix contains 3^3 variables. There are approaches implemented in TDEP that reduce the number of force constants that needs to be calculated. One such method is to use constraints based on crystal symmetry. The symmetry constraints result in Φ containing a combination of symmetry equivalent (reducible) and inequivalent (irreducible) components, effectively reducing the number of unknown variables. Another method affecting the number of force constants is the cutoff radius for atomic interactions, r_c . r_c defines a sphere within which two- or three-body interactions are considered. Outside the sphere, the interactions are set to zero.

Fig. 2.8 illustrates two third-order force constants as blue and red triangles. The red triangle is confined by a smaller r_c , and atoms i , j , and k are displaced from equilibrium, indicated by \mathbf{u}_i , \mathbf{u}_j , and \mathbf{u}_k . The blue triangle is bound by a larger r_c . This force constant would not be considered with r_c corresponding to the red circle. This illustrates how increasing r_c increases the number of force constants considered if more non-symmetry equivalent force constants are within r_c .

An example of how the number of irreducible force constants increases for increasing r_c is shown in Fig. 2.9. The results are calculated for a 432-atom supercell of orthorhombic ZnSb. This compound has fairly low symmetry when compared to e.g. cubic compounds. The number of force constants increases as a function of the cutoff radius, increasing r_c from 4 Å to 5 Å increases the number of force constants from 423 to 2772. Between 3.2 and 3.8 Å there is a plateau where the number of constants does not increase. In this region no new irreducible force constants are introduced; only atoms that correspond to symmetry equivalent force constants are being included, or alternatively, no new atoms are within the increasing r_c . An appropriate r_c can be obtained by convergence testing unless all possible interactions in the simulation cell are taken into consideration. Its value will depend on which material property is to be calculated and the desired accuracy. Within the TDEP framework, r_c can be set to different values for second- and third-order force constants.

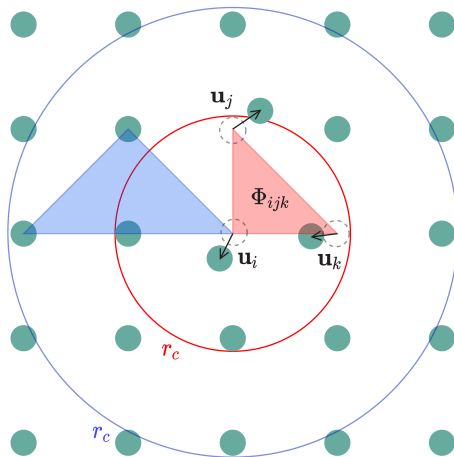


Figure 2.8: Illustration of third-order interatomic force constants. The green spheres indicate atoms, while the red and blue shaded triangles show two different three-atom interactions, each within a cutoff radius indicated by the red and blue circles. For the red triangle, atoms are displaced from their equilibrium positions. The illustration is made with inspiration from Ref. [89].

The total force on an atom is a result of all possible n -body interactions, while for the model the force on the atom depends on two- and three-body interactions that involve the atom. The model force on atom i in direction α can be expressed

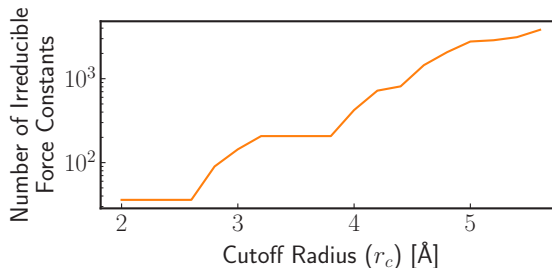


Figure 2.9: Number of irreducible force constants for orthorhombic ZnSb as a function of r_c . The horizontal axis indicates r_c for third-order force constants while the vertical axis shows the number of irreducible force constants.

as

$$f_{i\alpha}^M = \sum_j \sum_\beta \Phi_{ij}^{\alpha\beta} u_j^\beta + \frac{1}{2} \sum_{jk} \sum_{\beta\gamma} \Phi_{ijk}^{\alpha\beta\gamma} u_j^\beta u_k^\gamma. \quad (2.30)$$

All forces for the atoms in the cell can be grouped into the force vector \mathbf{F}_M . TDEP works by matching the forces of the model, \mathbf{F}_M , and the forces calculated with DFT, \mathbf{F}_{DFT} . By minimizing the difference between \mathbf{F}_M and \mathbf{F}_{DFT} we can obtain force constants that best represent the system. The minimization problem is expressed as

$$\min_{\Phi} \Delta \mathbf{F} = \frac{1}{N_c} \sum_{i=1}^{N_c} |\mathbf{F}_{\text{DFT}}^i - \mathbf{F}_M^i|, \quad (2.31)$$

where the sum goes over N_c configurations and the force difference is minimized using a least-squares solution.

Uncorrelated thermally excited configurations are desired as supercells that are used as input to DFT calculations to obtain \mathbf{F}_{DFT} . These configurations should be a diverse representation of the phase space. Ab initio molecular dynamics is often used to obtain realistic configurations at the desired temperature. A challenge with this approach is that many configurations (one from each time step) in the molecular dynamics run are correlated with each other. It is therefore useful to sample the molecular dynamics run, e.g. by only considering every 5th or 10th step. Thus, only a fraction of the calculations is useful as input to TDEP. With molecular dynamics, the system must also be equilibrated at the desired temperature before sampling.

An alternative to molecular dynamics is to generate configurations with a stochastic sampling approach [91]. This is implemented in the TDEP code [87, 89]. The approach limits computational expenses as no configurations need to be discarded. To create the configurations the instantaneous atomic positions, $u_{i\alpha}$, and velocities, $\dot{u}_{i\alpha}$, of atom i in a cell of N_a atoms is initiated by [92, 87]

$$u_{i\alpha} = \sum_{s=1}^{3N_a} \epsilon_{i\alpha}^s \langle A_i^s \rangle \sqrt{-2 \ln(\zeta_1)} \sin 2\pi\zeta_2 \quad (2.32)$$

$$\dot{u}_{i\alpha} = \sum_{s=1}^{3N_a} \omega_s \epsilon_{i\alpha}^s \langle A_i^s \rangle \sqrt{-2 \ln(\zeta_1)} \cos 2\pi\zeta_2. \quad (2.33)$$

Here, ω_s^2 is the eigenvalue of mode s , $\epsilon_{i\alpha}^s$ is the eigenvector of mode s , ζ are randomly drawn from a uniform distribution $(0, 1)$, and $\langle A_i^s \rangle$ is a thermally averaged amplitude in the classical limit. Details on how to approximate ω_s^2 and $\epsilon_{i\alpha}^s$ are discussed later. $\langle A_i^s \rangle$ can be obtained from [92],

$$\langle A_i^s \rangle \approx \frac{1}{\omega_s} \sqrt{\frac{k_B T}{m_i}}, \quad (2.34)$$

with T being the temperature and m_i the atomic mass. A requirement is that the positions and velocities should correspond to a canonical ensemble. If second-order force constants are available, $u_{i\alpha}$, $\dot{u}_{i\alpha}$, $\langle A_i^s \rangle$, and $\epsilon_{i\alpha}^s$ can be calculated straightforwardly by solving the equations of motion for the supercell. If second-order force constants are not available, an approximation can be made. These force constants can be sufficient for making uncorrelated configurations while being computationally efficient to obtain. A pair potential is the starting point, considering atom i and j at distance r_{ij} [87],

$$\frac{\partial U(r)}{\partial r_{ij}} = 0 \quad (2.35)$$

$$\frac{\partial^2 U(r)}{\partial r_{ij}^2} = \eta / r_{ij}^4. \quad (2.36)$$

For the equilibrium lattice, the first derivative is zero, the second derivative is positive and only substantial for small r_{ij} , and η is a constant. In this pair potential, the second-order force constants can be calculated analytically,

$$\Phi_{ij}(\mathbf{r}) = -\frac{\eta}{r^6} \begin{pmatrix} r_x r_x & r_x r_y & r_x r_z \\ r_x r_y & r_y r_y & r_y r_z \\ r_x r_z & r_y r_z & r_z r_z \end{pmatrix}, \quad (2.37)$$

with \mathbf{r} being the vector between atoms i and j . The parameter η tunes the force constants such that the phonons of the systems are within the correct frequency range. η can be obtained from the relationship between the zero-point energy of the phonons and the Debye model [87],

$$\frac{1}{N_a} \sum_i \frac{\hbar \omega_i(\eta)}{2} = \frac{9 k_B \theta_D}{8}. \quad (2.38)$$

To obtain η the Debye temperature, θ_D , must be calculated. This can be done using the following expression [48, 49],

2. Materials and Methods

$$\theta_D = \frac{h}{k_B} \left[\frac{3\rho}{4\pi} \right]^{1/3} v_a n^{-1/3}, \quad (2.39)$$

where ρ is the number of atoms per unit volume, n is the number of atoms in the primitive cell, and v_a is the average phonon velocity. v_a can be calculated with DFT, e.g. by using the elastic tensor from finite differences. Obtaining a set of thermally excited configurations can be done as follows: 1) the primitive cell is relaxed to obtain the equilibrium positions, 2) the elastic tensor is calculated, 3) an appropriately sized supercell is constructed from the relaxed cell, and 4) thermally excited supercell configurations are computed based on atomic displacements and velocities calculated with Eq. (2.35) and Eq. (2.36).

To obtain effective second- and third-order force constants at a specific temperature, T , we need configurations from an ensemble at T . The temperature is set in the amplitude, Eq. (2.34). Fig. 2.10 shows the absolute difference between atomic positions from a canonical ensemble, \mathbf{r}_{can} , and at equilibrium from DFT, \mathbf{r}_{equ} . The temperature of the canonical ensemble is set to 100 and 300 K. All atoms in configurations from the canonical ensemble are displaced. The largest position difference for the configuration at 100 K is $\sim 0.15 \text{ \AA}$ and at 300 K it is $\sim 0.15 \text{ \AA}$. The average absolute difference at 100 K is 0.056 \AA and at 300 K it is 0.085 \AA .

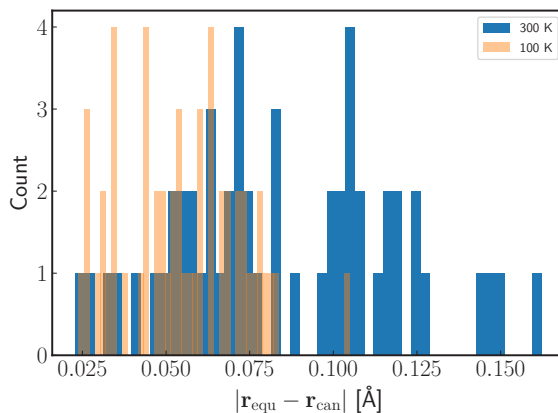


Figure 2.10: Histogram showing the displacements of atoms from equilibrium positions for supercell configurations with 108 atoms at 100 K and 300 K.

It should be mentioned that while the TDEP method is an effective way to obtain second- and third-order force constants, significant computational resources are required. In the TDEP manual, it is mentioned as a rule of thumb that convergence testing can be carried out starting at around 10 equations per irreducible force constant [90]. Each supercell configuration with forces calculated with DFT provides $3N_a$ equations, where N_a is the number of atoms. For a system with 128 atoms and 500 irreducible force constants, ~ 15 configurations

can be used as a starting point. Calculating the forces for a configuration with DFT roughly takes ~ 5 hours with 40 CPUs – 200 CPU-hours – for cubic compounds studied in this thesis. Obtaining the force constants using 15 configurations then requires ~ 3000 CPU-hours of computer time. Typically more than 15 configurations are necessary to obtain accurate third-order force constants, further increasing the computational expense. The computational expense also increases for systems with low symmetry, as the number of irreducible force constants increases. Calculating the third-order force constants – which are needed for the LTC – for thousands or tens of thousands of compounds is therefore a formidable task.

2.3.2 Relevance for the Articles

Article I:

In this article the LTC for 122 half-Heusler compounds is calculated based on DFT using the Vienna Ab initio Simulation Package (VASP) [93–95] and the TDEP method [83, 85]. The effective force constants are extracted at 300 K. The intrinsic LTC is considered with three-phonon scattering and scattering from isotopes using the natural distribution of isotopes. In this article, a machine learning model is trained where the features include the Debye temperature, Eq. (2.39) and LTC calculated with Slack’s model, Eq. (2.26).

Article II:

Here the LTC of the 122 half-Heusler compounds is calculated including scattering from isovalent sublattice substitutions and grain boundaries according to Eq. (2.24) and Eq. (2.25). The substitution concentration is set to 10 %, and the grain sizes are set to 100 and 200 nm. For each compound, 12 LTC calculations are done: intrinsic LTC, LTC with substitutions on the X -, Y -, and Z -sites, LTC with 100 and 200 nm grain boundaries, and LTC with substitutions on the X -, Y -, and Z -sites in addition to 100 and 200 nm grain boundaries.

Article III:

In this work, a machine learning method is trained on the intrinsic LTC of 268 compounds calculated with DFT and TDEP. Force constants are extracted at 300 K. 122 of these are the previously mentioned half-Heuslers, while 146 are cubic compounds from the Materials Project database [7]. The phonon dispersions, lifetimes, density of states, and spectral LTC for Na_2TlSb and Ca_3AsBr_3 are analyzed to identify possible origins of the low LTC.

Article IV:

In this article, eight compounds are found to have emerging electronic band gaps at the hybrid functional level. The LTC for these compounds is calculated using TDEP and three compounds, MgSc_2Hg , Li_2CaSi , and Ba_2HgPb , are found to have LTC less than 5 W/Km at 300 K.

Chapter 3

Results and Discussion

This chapter presents results and discussion on the topic of identifying low lattice thermal conductivity (LTC) compounds. Results that did not make it into the articles are also included. The first section of the chapter serves as an introduction to machine learning (ML) and features for material property predictions. The next part elaborates on ML for predicting the LTC, and possible avenues for choosing a representative training set, so-called active sample selection. The following sections concern LTC calculations. First is a discussion on how erroneous force constants affect calculated LTC, and second is a study on mass-disorder and grain-boundary scattering for reducing the LTC of AlVFe_2 . This is followed by a comparison of calculated LTC and thermal conductivities from experiments. Lastly, the chapter outlines research challenges and proposes new ideas for future research where the discussions are tied to the articles at the end of the thesis.

3.1 Machine Learning in Material Informatics

Machine learning (ML) is an umbrella term, along with "artificial intelligence", that encapsulates a large number of methods that *learn* something from data. The concept of model learning is close to how humans learn. Learning to ride a bicycle can be viewed as repeating similar movement patterns (getting up on the bike, pushing the pedals) to maintain balance. In ML terms this could be viewed as identifying the actions that contribute to and the actions that counteract balancing on the bike. Kids (typically) learn not to be rude to other kids. Often this comes through trial and error where the kids learn that their actions can inflict distress on others and/or trouble for themselves. From an ML point of view this could be thought of as how a model is "punished" when doing bad (making inaccurate predictions), and "rewarded" when doing good (making accurate predictions). In ML we use a set of *features* and a *model* that are used to predict a *target*. As an example, we can attempt to find features that can be used to predict how much ice cream will be sold by an ice cream stand on a given day. Features could be e.g. the temperature, whether it is sunny, and if it is a weekend. A warm, sunny Sunday could yield more sales than a cold, rainy Monday. Both features and target could be *classes*. The temperature is continuous, while "is it sunny" is binary and thus a class. It is also interesting to note how the features relate to the target. Would yesterday's stock price be a good feature for predicting ice cream sales? Several ML challenges must be resolved when modeling different problems. Which ML method should we choose and at what complexity level is desired? What features should be used in training? This section presents examples of ML applied in material science. Two

3. Results and Discussion

ML cases are explored: predicting cell volume from density functional theory (DFT) and LTC at different temperatures. This is followed by a comparison of active sampling methods.

3.1.1 Introduction: Predicting DFT Volume

Applying tools and ideas from informatics in materials science – materials informatics – has become a hot topic in recent years. The advent of materials databases that gather experimental and calculated data in a structured manner enables the use of statistical methods and ML [96–98]. As an introductory example take the simple linear regression model. The model can be expressed as

$$\hat{y}_i = C + \sum_{j=1}^p a_j x_i^j. \quad (3.1)$$

Here, C is a constant term, \hat{y}_i is the prediction for a sample i , and y_i is the target value. The target is the "true" value and is sometimes labeled "measurement". Each of the p features, x_i^j , are assigned a weight, a_j . In material informatics, the features are often labeled "material descriptors" or simply "descriptors". The model training is done by obtaining C and a_j that minimize the error between the predicted \hat{y}_i and target, y_i . The error can be computed as the residual sum of squares for N samples, $\sum_{i=1}^N (y_i - \hat{y}_i)^2$.

As a case study, we use an example from materials science; predicting the DFT relaxed volume per atom of a compound based on atomic properties. $N = 41$ compounds with two atoms in the primitive cell are used with four features, the average electronegativity, $\langle \chi \rangle$, average covalent radius, $\langle r \rangle$, standard deviation of the electronegativity, σ_χ , and the standard deviation of the covalent radii, σ_r . By minimizing the residual sum of squares the following model expression is obtained,

$$V_{\text{model}} = -9.0 - 5.6 \langle \chi \rangle + 35.7 \langle r \rangle + 5.4\sigma_\chi - 17.1\sigma_r. \quad (3.2)$$

χ is the Pauling electronegativity and the unit of r is Å. The predictions of the model using Eq. (3.2) are shown in Fig. 3.1. The average relative error of using Eq. (3.2) is 13 %.

The relationship between the features and the volume per atom can be interpreted. A large $\langle r \rangle$ corresponds to a large effective size of the atoms and hence a larger V_{model} . When r is different between the two atoms in a compound σ_r becomes large and V_{model} is smaller. It is also interesting to note that this model can extrapolate beyond material volumes seen in training. For example, a two-atom system containing two fluorine atoms has a negative V_{model} from Eq. (3.2).

The problem of extrapolating to a negative volume can be solved by manipulating the problem formulation. With the natural logarithm of the volume and the features, the following expression is obtained,

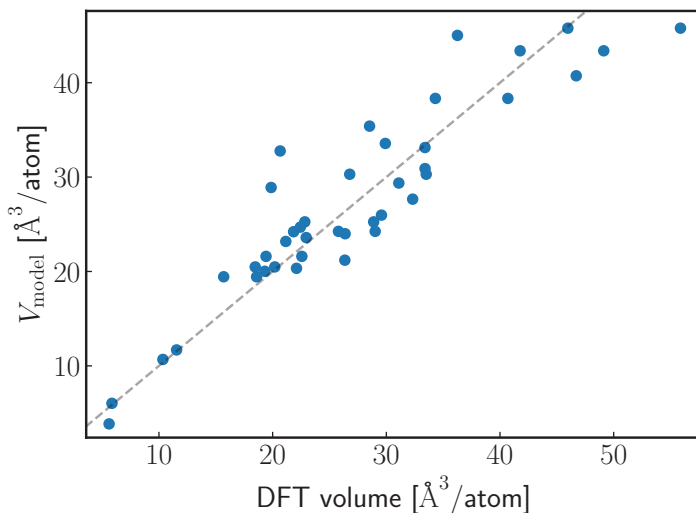


Figure 3.1: Linear regression model for predicting the compound volume per atom. The horizontal axis shows the volume per atom calculated with DFT, and the vertical axis shows the predicted volume, V_{model} , calculated with Eq. (3.2).

$$\log(V_{\text{model},\log}) = C + a_1 \log(\langle \chi \rangle) + a_2 \log(\langle r \rangle) + a_3 \log(\sigma_\chi) + a_4 \log(\sigma_r). \quad (3.3)$$

Training of the linear regression model yields the following expression after taking the exponent on both sides,

$$V_{\text{model},\log} = e^{2.3} \langle \chi \rangle^{0.04} \langle r \rangle^{2.3} \sigma_\chi^{-0.06} \sigma_r^{-0.07}, \quad (3.4)$$

This model can not extrapolate to negative values and performs similarly to the previous regression model. The predictions are shown in Fig. 3.2 and the average relative error is 12 %. The exponents of $\langle \chi \rangle$, σ_χ , and σ_r are relatively small. Small changes in these features make less difference in the predictions. Fig. 3.2 also shows the volume predicted by a model that is trained using only one feature, $\langle r \rangle$, resulting in the following expression,

$$\hat{V}_{\text{model},\log} = e^{2.6} \langle r \rangle^{2.1}. \quad (3.5)$$

The constant term increases slightly while the exponent of $\langle r \rangle$ decreases. The error is 17 %, a slight increase over the previous models.

The linear regression is straightforward to use and the simple expressions obtained here for the DFT compound volume can be useful, e.g. as a starting point before DFT relaxations. It could be extra beneficial for reducing the computational expense if a large number of compounds are to be relaxed in

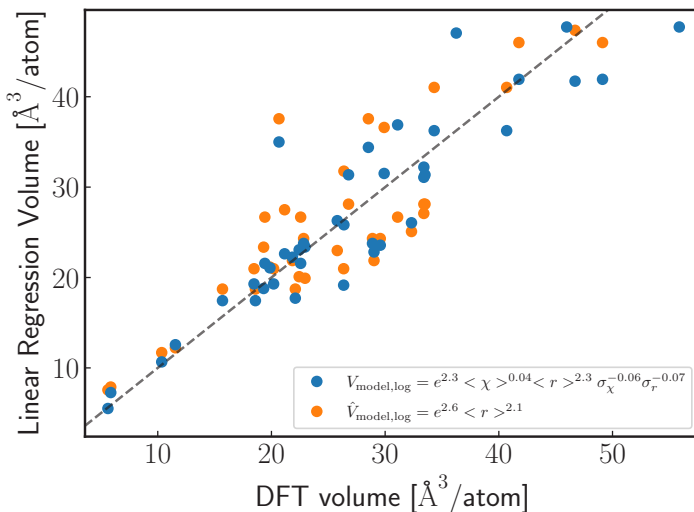


Figure 3.2: Linear regression model for predicting the volume per atom. The horizontal axis shows the volume per atom calculated with DFT, and the vertical axis shows the predicted volume calculated with Eq. (3.4) (blue disks) and Eq. (3.5) (orange disks).

a high-throughput fashion. It should be noted that models specialized for compound sub-classes can achieve better performance than the models obtained here. A model made by Miyazaki et al. had an error of $\sim 5\%$ between computed and predicted lattice parameters for half-Heuslers [99].

Absorbing the details of ML and materials informatics, such as feature engineering, model selection, validation schemes, data selection, and so on can be a challenge. This blog post from the Citrine web page [100] and the links within give a short introduction to material informatics. The work by Wang et al. [101] provides broad guidelines for the field, and the authors include ML examples in Jupyter notebook format. The advantage of notebooks is that they are easily adapted to different data. The work by Seko et al. [102] tackles one of the main challenges in ML for compound properties; how to generate structured features from compound sets.

3.1.2 Features for Machine Learning in Materials Science

Features in material informatics are descriptors that represent a material in a structured way. Features should have a general form such that the features can be obtained for all materials studied. E.g. inter-layer distance can be a useful feature for layered materials while it is hard to define such a distance for non-layered materials. Features for specific compound classes can enhance model performance. Miyazaki et al. studied the half-Heusler compounds using

ML and features tailored to the composition for predicting the lattice thermal conductivity and volume [99]. The half-Heuslers have three atoms in the primitive cell, XYZ , and Miyazaki et al. found that the features $(r_X + r_Y + r_Z)/3 - r_X$ and $(m_X + m_Y + m_Z)/3 - m_X$, where r was the atomic radius and m was the atomic mass, were important features. Specialized features such as the absolute difference of atomic radius of elements in binary compounds, $|r_A - r_B|$ and the sum, $r_A + r_B$, have been used to determine bonding types in compounds [103]. These features enabled the identification of compositions AB that had large differences in energy when comparing zinc blende and rock salt crystal structures.

To make general features based on compositions one can start with a vector, $l = [a_1, \dots, a_n]$, where a is a descriptor (such as atomic radius or mass), and n is the number of atoms in the primitive cell. A possible starting point for effectively obtaining a is the Mendeleev python package [104], which contains different atomic descriptors. Matminer [105] can also be used to obtain a from descriptors based on the atoms or the structure. The function $F(l)$ enables l to be converted into a feature and can handle l of any length larger than zero. $F(l)$ can take many forms, such as the minimum, maximum, average, median, standard deviation, number of elements, or number of unique elements. The point of $F(l)$ is that it should be able to map l of any length to a value. It is also possible to take information from the crystal structure itself and use it to construct features, e.g. the space group or average coordination number. The Voronoi structure has been used to calculate features related to symmetry and structure in a general fashion [106, 107]. The Voronoi structure for a lattice point is defined by planes perpendicular to the vector connecting the point with its closest neighboring lattice points at the halfway point of the vector. Fig. 3.3 shows an example of a 2D Voronoi structure. The number of faces and shape of the faces of the Voronoi structure capture part of the local environment in the structure. For a given compound, a for an atom could for example be given by the number of faces or area of all faces.

The previously discussed features are based on the atoms and the lattice. Another option is to calculate features with DFT. Such features are useful when the calculation of the features is much less computationally expensive when compared to obtaining the target property. Two examples are the relaxed volume per atom and the bulk modulus. The role of these features when using ML to predict the LTC is explored in Article I. The DFT computed electronic band structure provides features, e.g. the electron effective mass and band gap. Phonon dispersions can be somewhat more computationally expensive to obtain, while they provide features linked to lattice dynamics, such as phonon group velocities, maximum frequencies, effective spring constants, and first-moment frequencies [56].

Overfitting often occurs in ML and may worsen if the number of features is large. Using a large number of features can also lead to runtime or memory issues. Feature selection schemes enter as useful tools for reducing the number of features. These schemes work by removing redundant or correlated features. There are different types of schemes providing good feature subsets, many of which have been employed with ML models predicting the LTC [99, 108–110].

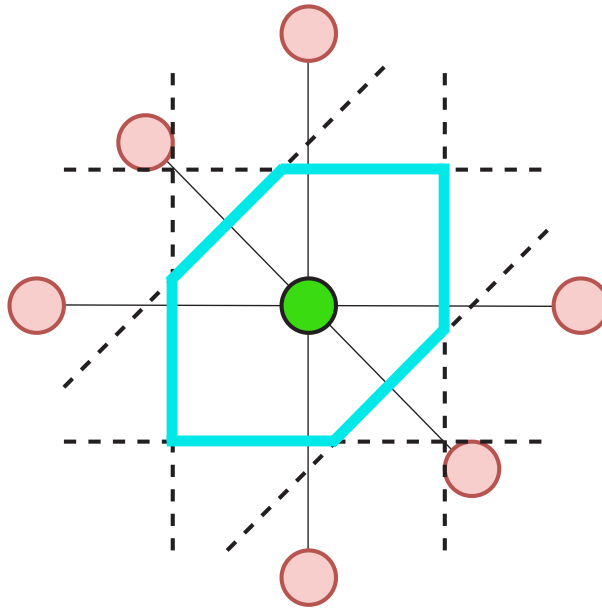


Figure 3.3: Voronoi structure illustration (turquoise) for the central atom (green).

In Article I we use a scheme called exhaustive feature selection. This scheme is computationally costly and brute-forces the best feature subset by evaluating all possible feature subsets. Article III employs an iterative scheme, sequential feature selection, that identifies a feature subset by observing the performance drop when feature values are randomly shuffled.

Domain knowledge is crucial for constructing the initial feature set. The bulk modulus can for example be related to atomic bonding and phonon transport, and the number of atoms in the unit cell can affect the phonon dispersions and the fraction of acoustic branches [39]. Knowing which features are more or less relevant for the target property can also reduce feature redundancy, as less relevant features can be removed.

3.1.3 Machine Learning the Lattice Thermal Conductivity: Which Model is Best?

Using ML to predict the lattice thermal conductivity (LTC) is a core topic in this thesis. Evaluating the ML performance of different models is important for attaining high accuracy. The works by Zhu et al. [110] and Miyazaki et al. [99] contain performance comparisons of popular ML models. Different ML models have different strengths and weaknesses, making it challenging to identify the optimal one. Examples of models used for predicting the LTC are random forests regression (RFR) [110–113], boosted trees methods [99, 112], Gaussian process regression (GPR) [108, 114], and neural networks [110, 115].

Jaafreh et al. [112] trained an ML method using a dataset of 119 compounds with LTC at different temperatures. The LTC was obtained based on DFT calculations, Phonopy [79], and Phono3py [80]. They showed that good performance can be achieved when predicting the LTC at different temperatures using the popular RFR model. Fig. 3.4 shows a decision tree, the core component in RFR. The decision tree starts with a sample (compound) at the top and propagates it through the logical splits (rectangles) until a leaf node is reached (circles). In this decision tree the compound can be evaluated on four features: The average mass, m_{avg} ; the volume, V ; the average covalent radius, r_{avg} ; and the number of elements, N . RFR combines many different decision trees in a randomized fashion. Each decision tree is based on random subsets of features and compounds, resulting in a variety of trees that make up the "forest". When a sample is propagated through all decision trees of the forest, its predicted value is the average across predictions for individual decision trees, $\text{LTC}_{\text{avg}} = (1/N) \sum_{i=1}^N \text{LTC}_i$, where the sum goes over all decision trees. Since the leaf nodes are LTC from compounds used for training, LTC_{avg} is bound by the lowest and highest LTC_i in the training set.

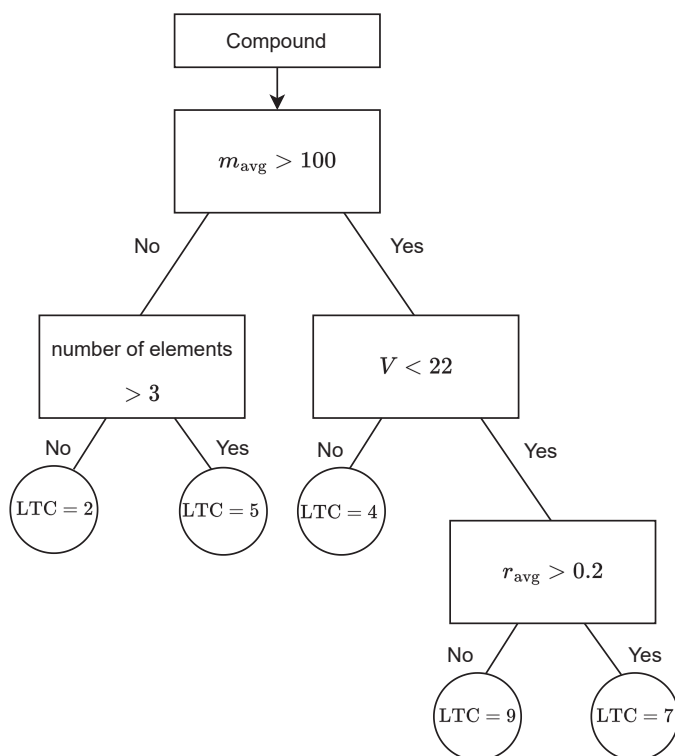


Figure 3.4: Illustration of a decision tree.

In the following, we use different ML models and assess the performance when predicting the LTC. The LTC is calculated with the temperature-dependent

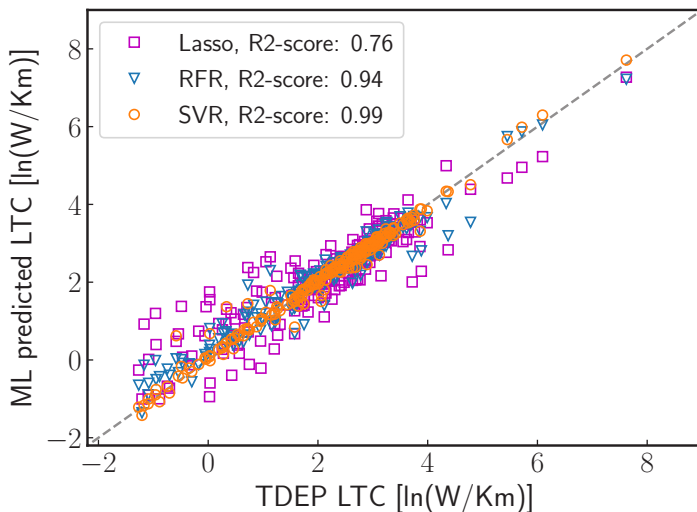


Figure 3.5: Predictions of the LTC using RFR, SVR, and Lasso.

effective potential (TDEP) method. The dataset consists of 220 compounds with two, three, and four atoms in the primitive cell. Five temperatures are used, 200 K, 300 K, 400 K, 500 K, and 600 K, making a total of $220 \times 5 = 1100$ samples. As features, the average, standard deviation, median, minimum, maximum, and ratio of minimum to maximum of the following properties of the atoms in the primitive cell are used: electronegativity, electron affinity, dipole polarizability, valence, covalent radius, and mass. The volume per atom, number of atoms in the primitive cell, number of unique atoms in the primitive cell, and the temperature are also included. In the following RFR is compared to two other ML models, support vector regression (SVR) and Lasso. The SVR model is non-linear. In simplified terms, it makes predictions based on how close an unseen compound is in feature space to other compounds. The Lasso model is linear, but different from the simpler linear regression, it minimizes the sum of the weights, a_j in Eq. (3.1), with so-called regularization. The performance is computed as an average over a 5-fold cross-validation (CV) where hyperparameters are optimized using a randomized grid search.

Fig. 3.5 shows the predictions made by the three models from the 5-fold CV compared to the TDEP LTC. The R2-score is used to evaluate prediction accuracy. It is defined as $R2\text{-score} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$, where y_i is the calculated LTC for sample i , \hat{y}_i is the ML predicted LTC, and \bar{y} is the average of the calculated LTC. The R2-score for Lasso, RFR, and SVR, is 0.76, 0.94, and 0.99, respectively. The R2-scores for the RFR and SVR models are comparable. While the R2-score using the Lasso model is substantially lower, training this model is much faster than training the RFR and SVR models.

Six compounds are withheld from the training and validation process as a test

set. These compounds are LaPtSb, TiNiPb, CuCl, RbBr, Li₂InBi, and ZrSiRu₂. The predictions for the compounds are shown in Fig. 3.6. The R²-scores are -0.19, 0.72, and 0.22 for Lasso, RFR, and SVR, respectively. The RFR and SVR models are the least accurate for CuCl. If CuCl is removed from the test set, the R²-scores are -0.30, 0.89, and 0.91 for Lasso, RFR, and SVR.

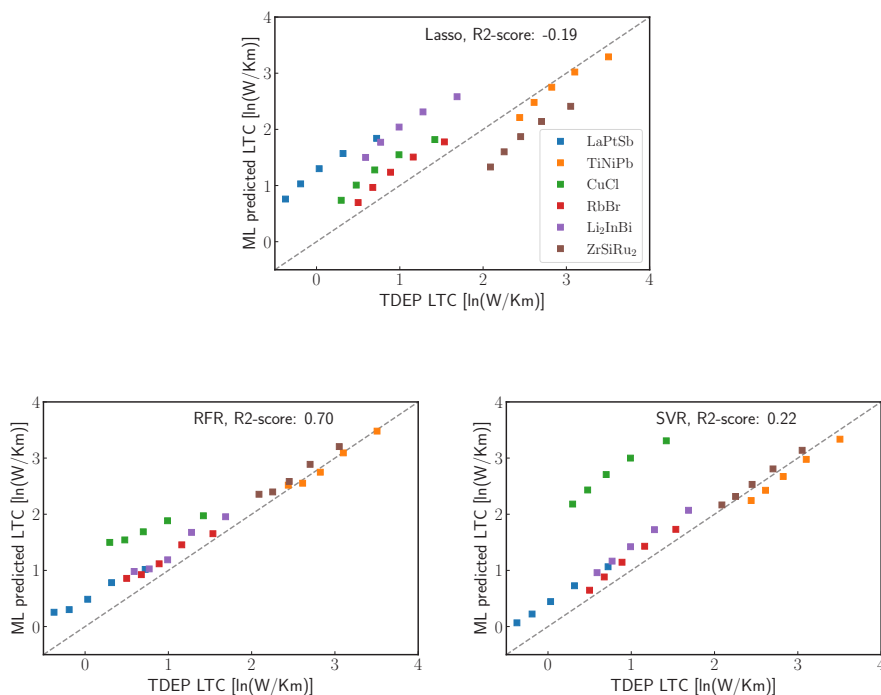


Figure 3.6: ML LTC predictions on the test set using Lasso (upper panel), RFR (lower left panel), and SVR (lower right panel). For each compound, the LTC is predicted at 200, 300, 400, 500, and 600 K.

The features used in these ML models come from tabulated data or computationally inexpensive DFT calculations. There have been proposed several descriptors linked to anharmonicity and low LTC, such as first frequency moments [56], three-phonon scattering phase space [115], and statistical anharmonicity measures [116]. Using such descriptors as features in ML models can increase performance, and is viable if the computational expense of obtaining the descriptors is sufficiently low. A descriptor based on second-order force constants could be approximately an order of magnitude less computationally expensive to obtain than the LTC itself [117].

The RFR and SVR models performed worst when predicting the LTC for CuCl. If no training compounds exhibit the same properties as CuCl it could be

3. Results and Discussion

challenging to recognize it as a low LTC compound. Ensuring a representative and diverse training set is investigated in Article I and Article III. The following subsection provides a discussion of the concept of active sample selection.

3.1.4 Active Sample Selection

Active sample selection – also called *active learning* – is important when the training data is not necessarily representative of the data set we want to explore. Small training sets often require active sampling, as the size alone does not enable the sets to be representative. When predicting the LTC, obtaining training sets becomes a challenge because computing the LTC for training is computationally expensive. Sample diversity is also important with larger training sets if the unseen data is non-uniform and distributed. Fig. 3.7 illustrates how the structure of the data can affect how its sampled, where the sampled data points are green crosses. The data on the left is connected and uniform, while on the right it is separated. One of the sub-regions of the right data is left un-sampled as a result.

Similar techniques to active learning have also been employed in the DFT community. VASP [93–95] contains a method for on-the-fly ML force fields where DFT calculations are done for configurations with high ML uncertainty [118]. A Bayesian linear regression model provides energy predictions and uncertainties. Adversarial attacks have been used to expand training domains for neural network interatomic potentials [119]. New configurations to be calculated with first principles methods are found iteratively by evaluating the difference in energy predictions of neural networks trained on different data subsets. Training cluster-expansion models is computationally expensive, requiring large sets of configurations evaluated with DFT. It has been found that a group of models (an ensemble) trained on overlapping subsets of the data can find configurations that should enter the training set [120], effectively reducing the number of DFT calculations needed. A review of active sampling strategies, utility functions, and challenges in materials science can be found in Ref. [121].

Principal component analysis (PCA) can be used for active sample selection, as explored in Article I. PCA acts as a dimensionality reduction, taking correlated features into account. Principal components are made by projecting points in feature space onto a new set of axes that maximize the variance. An example of the construction of a principal component (PC) is shown in Fig. 3.8. The original space is spanned by two features, X_1 and X_2 . The blue points indicate different compounds in the space. The projections of the blue points onto the line are shown as green points, where the line is computed such that the green points have their variance maximized. This defines the first PC. On the axis defined by the orange line, the projections are scalars because of the dimensionality reduction. This concept can be expanded for more features and perpendicular first, second, third, and so on, PCs can be constructed. There are several options when it comes to using PCA for active sampling, including choosing which or how many PCs should be used to span the space. One approach is to iteratively include the compound with the largest Euclidean distance from any compound

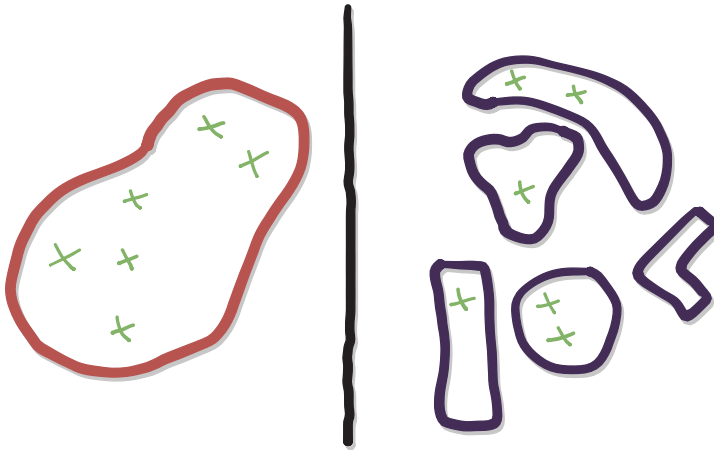


Figure 3.7: Examples of distributed data. The red shape and purple shapes represent different data sets. The green crosses are sampled data points.

in the training set in the PC space. This approach, using the two first PCs, is used in Article I.

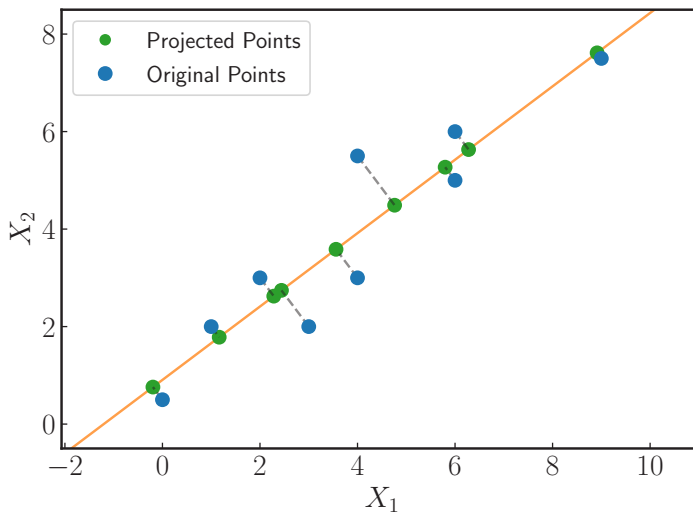


Figure 3.8: The horizontal and vertical axes show two different features (compound descriptors). The original data (blue points) are projected onto the orange line which corresponds to the first PC.

Gaussian process regression (GPR) is a probabilistic regression method and is used in Article III, both for doing predictions as well as active sampling. In

3. Results and Discussion

contrast to PCA, using GPR for active sampling requires a training set with target values, i.e. the calculated LTC. For GPR the covariance function, often called kernel, is central. It measures the similarity of different compounds. There is a variety of kernels available and the choice of a suitable kernel depends on the underlying structure of the data. A common kernel is the radial basis function (RBF) kernel,

$$k(x, x') = \exp \left[\frac{-(x - x')^2}{2\ell^2} \right]. \quad (3.6)$$

Here, ℓ is a characteristic length-scale parameter, and x and x' are variables representing two compounds. When x and x' are similar, $k(x, x')$ goes towards unity, and similarly, $k(x, x')$ goes towards zero for dissimilar x and x' . With n training compounds and a new compound x_* we define three matrices to contain the covariance functions between all samples [122],

$$K = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{pmatrix}, \quad (3.7)$$

$$K_* = [k(x_*, x_1), k(x_*, x_2), \cdots, k(x_*, x_n)], \quad (3.8)$$

$$K_{**} = k(x_*, x_*). \quad (3.9)$$

If \mathbf{y} is the vector of targets for the n compounds the probability of a prediction y_* given \mathbf{y} is $p(y_* | \mathbf{y})$. This probability corresponds to a Gaussian distribution,

$$y_* | \mathbf{y} \sim \mathcal{N}(K_* K^{-1} \mathbf{y}, K_{**} K^{-1} K_*^T). \quad (3.10)$$

The estimate for y_* is then the mean of the distribution,

$$\bar{y}_* = K_* K^{-1} \mathbf{y}. \quad (3.11)$$

It then follows that the uncertainty of the new prediction y_* can be obtained through the variance of the prediction,

$$\text{var}(y_*) = K_{**} - K_* K^{-1} K_*^T. \quad (3.12)$$

Fig. 3.9 illustrates GPR predictions and uncertainties where training is done on nine data points. The prediction variance increases in regions with few or no training data.

GPR-based active sampling can be done as follows: 1) train the model with available LTC, 2) do predictions on the unseen compounds, 3) find those with uncertainty over a predefined threshold, 4) calculate the LTC for these compounds, 5) include the new compounds in the training. It is also possible to do it differently by, 3) finding the compound with the highest uncertainty, 4) using the ML predicted LTC as the target value and including it in the training

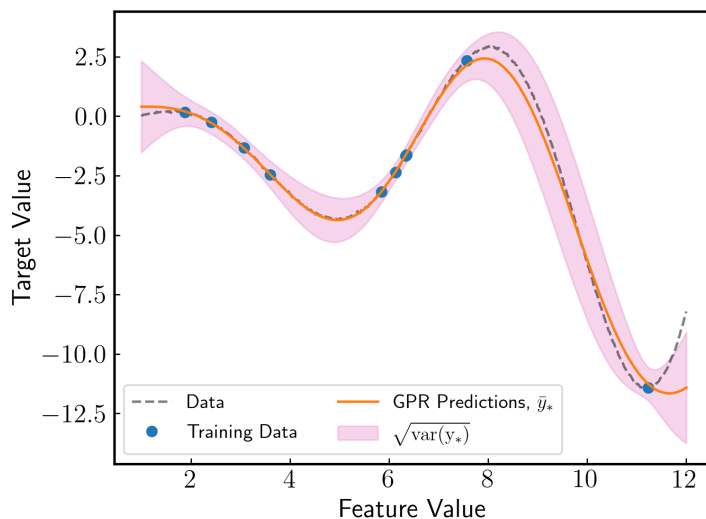


Figure 3.9: GPR illustration with predictions (orange line) and uncertainties (purple).

set, 5) repeating 4) until a suitable set of compounds is obtained and then calculate the LTC for the set. This approach is similar to that done in Article III. Both these approaches enable calculation of the LTC in bulk for the whole set of actively sampled compounds – an advantage when using high-performance computational clusters.

Fig. 3.10 shows a comparison of compounds chosen with active sampling using PCA and GPR. The scenario is the same as in Article I, with a test set of 35 compounds and a training pool of 87 compounds. Note that the five compounds with the lowest LTC are placed in the test set. The objective is to find compounds that should be moved from the test set to the training pool to increase the accuracy of predictions on the test set. The compound included with PCA-based sampling is the one in the test set with the largest Euclidean distance to any compounds in the training pool. The compound included with GPR-based sampling is the one in the test set with the highest GPR uncertainty. In the GPR sampling, the compounds selected are iteratively included in the model training. In the first iteration, both methods choose to include BaBiK. In the second iteration, PCA sampling includes CdPNa, while the GPR sampling includes LiZnSb. When five compounds are sampled four are the same across the two methods, BaBiK, CdPNa, LaPtSb, and GaNiNb. Only the PCA-based sampling includes LaRhTe, while only the GPR-based sampling includes LiZnSb.

Next is a case study with different sampling schemes and LTC predictions done by GPR models, similarly as done in Article III. Fig. 3.11 displays the Spearman rank correlation and RMSE between predicted LTC and TDEP LTC with models based on three sampling schemes: random, GPR STD, and PCA.

3. Results and Discussion

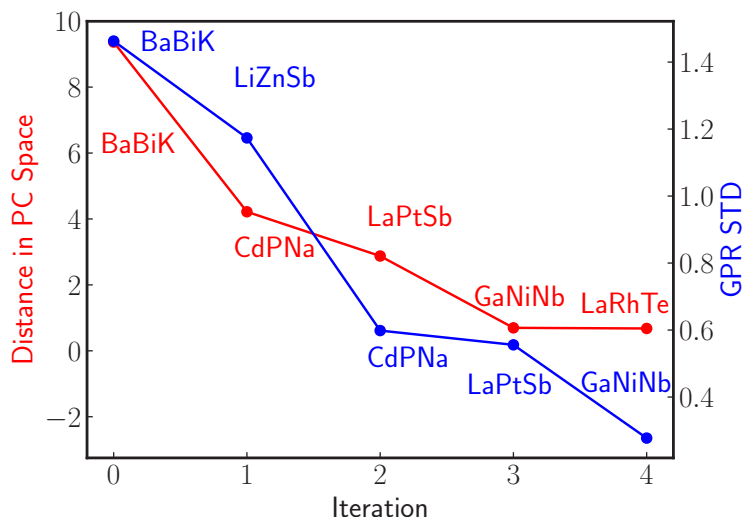


Figure 3.10: Illustration of compounds from active sampling using PCA and GPR STD. The left axis (red) shows the distance in PC space between the compound in the test set farthest from the training pool and its closest training pool neighbor. The right axis (blue) shows the GPR STD of the compound with the highest GPR STD in the test set.

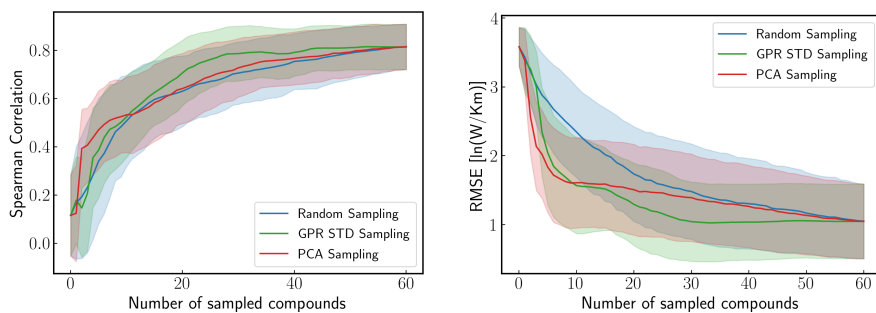


Figure 3.11: Spearman correlation (left panel) and RMSE for the three compounds with the lowest TDEP LTC (right panel) for predictions made on the test set (vertical axis) vs. number of compounds sampled to the training set (horizontal axis).

The RMSE is taken for the three compounds with the lowest TDEP LTC. The initial training set (0 on the horizontal axis) is made of 151 compounds with three atoms in the primitive cell. The pool that the models can sample from consists of 60 compounds with 2, 4, 5, and 6 atoms in the primitive cell. The test set is made of 27 compounds with the same number of atoms in the primitive cell as the pool. The compounds in the pool and test set are randomly shuffled 150 times to provide the standard deviation of plotted values. The Spearman rank correlation can be used to measure how the models differentiate low and high LTC compounds. It describes with a value $[-1, 1]$ how well two variables can be fitted by a monotonic decreasing, $[-1, 0)$, or increasing, $(0, 1]$, function. For less than 11 sampled compounds PCA sampling has the highest Spearman correlation and lowest RMSE. When more than 11 compounds are sampled the GPR STD-based model outperforms the others. With GPR STD-based sampling a correlation > 0.78 is achieved with 28 sampled compounds, PCA-based and random sampling need 45 and 47 compounds to achieve similar scores.

3.2 Lattice Thermal Conductivity Calculations

This section presents results and discussions concerning lattice thermal conductivity (LTC) calculated with the temperature-dependent effective potential (TDEP) method. First, the section outlines how noise added to third-order force constants affects calculated LTC. Second is a case study where mass-disorder and grain-boundary scattering is used to reduce the LTC of AlVFe_2 – a thermoelectric compound with high intrinsic LTC. Here it is also shown how different scattering mechanisms scatter phonons of different energies. Lastly, the section presents a comparison of calculated LTC and experimental values.

3.2.1 Lattice Thermal Conductivity with Noisy Force Constants

In the previous section, Article I, and Article III, the core objective is to predict the LTC directly using machine learning (ML). An alternative to this is to do ML for predicting the interatomic force constants (FCs) (or irreducible force constants (IFCs) that are mapped back to FCs) instead and then solve the phonon Boltzmann transport equation (BTE) to obtain the LTC. The computational expense of solving the phonon BTE is modest compared with the expense of doing supercell force calculations with DFT. The ML-FCs are from one point of view more flexible than a direct ML prediction of the LTC, as the ML-FCs can be used for other purposes, such as calculating phonon dispersions and Grüneisen parameters. Carrete et al. [111] used a PCA-based method to predict the IFCs. Obtaining the IFCs and the LTC only required a handful of DFT supercell calculations and the solution of the phonon BTE for each compound. Using this method they achieved higher accuracy than when directly predicting the LTC. There have also been other efforts made to reduce the computational cost of obtaining FCs, e.g. using compressive sensing [123] and ML [82, 124], which yielded accurate LTC. The goal of this subsection is not to predict the FCs

3. Results and Discussion

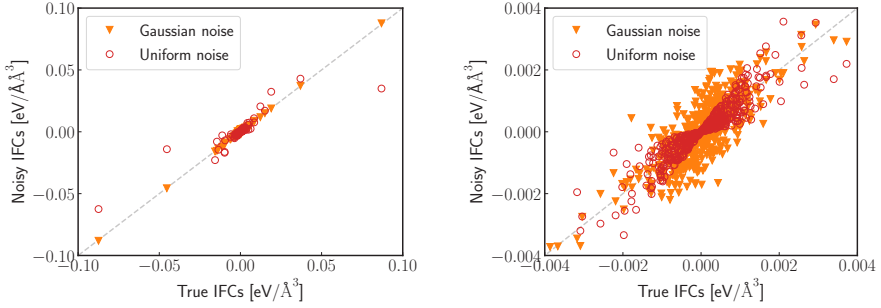


Figure 3.12: Values of the true IFCs are shown on the horizontal axis and noisy IFCs on the vertical axis. (a) shows all IFCs, while (b) shows IFCs close to zero.

with ML, but rather analyze what happens to the LTC when IFCs are slightly inaccurate, as one could expect from an ML model.

In the following, noise is introduced in the third-order IFCs for AlVFe_2 . The noisy IFCs are obtained by adding random noise drawn from a Gaussian distribution to each IFC. The Gaussian distribution has a mean of the average of the IFCs, $\langle \text{IFC} \rangle$, calculated with DFT and TDEP (labeled true IFCs), and standard deviation of $0.1 \cdot \sigma_{\text{IFC}}$, where σ_{IFC} is the standard deviation of the true IFCs. The i th noisy IFC is then $\text{IFC}_i^{\text{Gaussian}} = \text{IFC}_i + \epsilon$, where ϵ is drawn from the distribution. Fig. 3.12 shows a comparison of the true IFCs and the IFCs with Gaussian noise displayed as orange triangles. The average relative error of the noisy IFCs is $\sim 300\%$. For the large IFCs close to $\pm 0.9 \text{ eV}/\text{\AA}^3$ the absolute error of the noisy IFCs is low compared to the absolute values of the IFCs.

Another set of noisy IFCs is made by multiplying each IFC with a random number, N , drawn from a uniform distribution $0.25 \leq N \leq 1.75$, $\text{IFC}_i^{\text{Uniform}} = \text{IFC}_i \times N$. Fig. 3.12 displays these IFCs as red circles. The average relative error of the IFCs is $\sim 37\%$. This choice of noise has a noticeable effect on IFCs with high absolute value.

Changing the IFCs directly impacts the three-phonon scattering rate, Eq. (2.21). Fig. 3.13 shows the LTC calculated with the true IFCs, and an uncertainty region bound by the LTC calculated with the noisy IFCs. The second-order IFCs are left unchanged. Using IFCs with Gaussian noise results in an LTC that is $\sim 20\%$ lower than with true IFCs at 300 K. Using uniform noise the LTC is $\sim 20\%$ higher. The sum of the absolute value of the true IFCs is $0.75 \text{ eV}/\text{\AA}^3$, with Gaussian noise $0.86 \text{ eV}/\text{\AA}^3$, and uniform noise $0.67 \text{ eV}/\text{\AA}^3$. The three largest $\text{IFC}_i^{\text{Uniform}}$ have smaller absolute values than their true IFC counterparts. Decreasing (increasing) values of the third-order IFCs can result in lower (higher) three-phonon scattering and thus higher (lower) LTC.

The uncertainty is only based on two LTC calculations, one using $\text{IFC}^{\text{Gaussian}}$ and one using $\text{IFC}^{\text{Uniform}}$, and in this case, $\text{IFC}^{\text{Gaussian}}$ results in a lower LTC

and $\text{IFC}^{\text{Uniform}}$ results in a higher LTC. A different random seed can give noisy IFCs that differ from those used here. In the ideal case 10s or 100s of LTC calculations with $\text{IFC}^{\text{Uniform}}$ and $\text{IFC}^{\text{Gaussian}}$ using different random seeds should be performed to obtain an accurate uncertainty region.

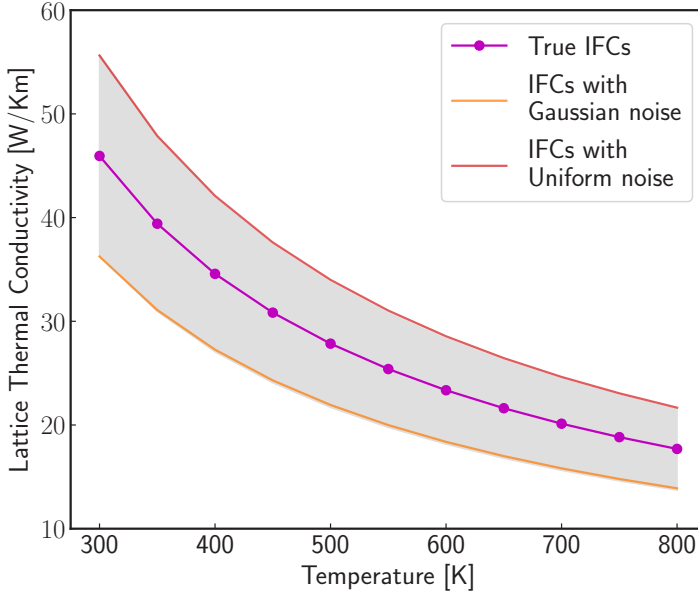


Figure 3.13: LTC as a function of temperature for AlVFe_2 using true IFCs (purple), with uncertainty based on calculations using IFCs with Gaussian noise and uniform noise.

To check if even larger errors in large IFCs can be acceptable the LTC is calculated with noisy IFCs using a uniform distribution as previously using an interval $0.15 \leq N \leq 1.85$. These IFCs have an average relative error of 40 %. Using these IFCs the LTC calculation fails to converge. Carrete et al. [111] found that calculations could fail to converge when using IFCs from Mg_2Si to calculate the LTC of other half-Heusler compounds. Correctly identifying and predicting large-value IFCs with ML can be a requirement to ensure the accuracy and convergence of LTC calculations.

3.2.2 Atomic Substitutions and Grain Boundaries Reduce Lattice Thermal Conductivity in Full-Heusler AlVFe_2

While exploring new thermoelectric compounds one can encounter candidates with favorable electronic properties but high LTC, resulting in a low ZT and poor thermoelectric performance. For these compounds, a viable option is to reduce the LTC by introducing additional phonon scattering mechanisms such as

3. Results and Discussion

alloying elements or grain boundaries (gbs) in manufacturing. The effect of such scattering mechanisms can be simulated with computational methods before the costly and difficult manufacturing, guiding the compound design for lowering the LTC.

AlVFe_2 is a cubic full-Heusler compound, XYZ_2 , as illustrated in Fig. 3.14. This thermoelectric compound has been recently studied, with experimental $zT \sim 0.1$ for $\text{Al}_{1-x}\text{Ge}_x\text{VFe}$ [125], ~ 0.2 for $\text{Al}_{1-x}\text{Ta}_x\text{VFe}$ [126], and ~ 0.3 for $\text{Al}_{1-y}\text{Si}_y\text{V}_{1-x}\text{Ta}_x\text{Fe}$ [127]. These values are low compared to high-performance thermoelectric compounds, but the interest in the compound could partially be attributed to the $zT \sim 5$ obtained for a metastable phase in thin-film $\text{AlV}_{0.8}\text{W}_{0.2}\text{Fe}_2$ [128], as well as its favorable electronic transport properties at room temperature [127]. In Article IV it was found that AlVFe_2 in the full-Heusler structure could achieve ZT of ~ 0.25 at 300 K with a fixed LTC of 4 W/Km using theoretical calculations. The LTC calculated with TDEP is much higher than 4 W/Km, it is ~ 46 W/Km at 300 K. It is therefore interesting to uncover whether the LTC of this compound comes close to 4 W/Km when including additional phonon scattering mechanisms in calculations. In the following case study, the phonon dispersions and LTC are calculated with TDEP, and partial sublattice substitutions and gbs are introduced to reduce the LTC of AlVFe_2 .

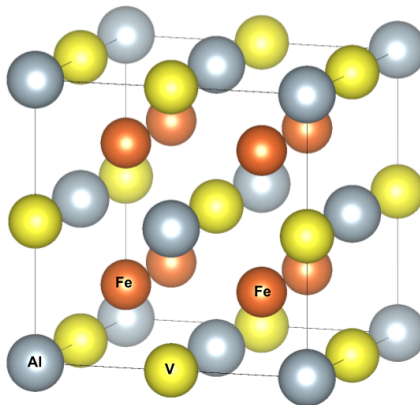


Figure 3.14: Crystal structure of AlVFe_2 .

The phonon dispersions and site-projected DOS in Fig 3.15 shows which atoms vibrate at different energies. There is a energy gap between low- and high-energy optical phonons. Gaps between phonon branches can prevent three-phonon scattering events [129]. This could partially explain why AlVFe_2 has a relatively high intrinsic LTC, although branch gaps have been found for low LTC compounds, e.g. for the half-Heuslers BaBiK and CdPNa [56]. Acoustic phonons have high group velocity for low energies and can contribute substantially to the LTC [130], as seen from Eq. (2.19). Suppressing acoustic phonons is thus a viable route for reducing the LTC. The Al atom is the lightest and has high

projected DOS for high-energy optical phonons. These phonons are above the gap in the phonon dispersions. The heavier V and Fe atoms have high DOS for middle energy phonons, in the range of 25-40 meV. There are both acoustic and optical phonons in this range below the dispersion gap. For low-energy vibrations, the Fe atom has the highest projects DOS.

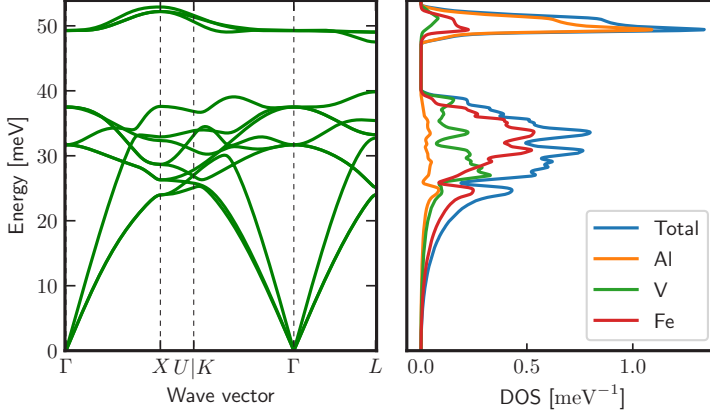


Figure 3.15: Phonon dispersion (left panel) and DOS (right panel) for AlVFe_2 . The projected DOS for the two equivalent Fe atoms are summed.

Partial substitution is done using the elements within the same group in the next period. These substitutions maintain the same number of valence electrons in the structure and are often labeled as isovalent or isoelectronic substitutions. Thus, Al is substituted by Ga, V by Nb, and Fe by Ru. Substituting with a heavier element can lead to a high mass variance parameter, g_i , Eq. (2.23), increasing the mass-disorder scattering rate, Eq.(2.24).

In the following we label the intrinsic LTC as κ_ℓ^{int} , with Al substituted by Ga, $\kappa_\ell^{\text{Al/Ga}}$, V substituted by Nb, $\kappa_\ell^{\text{V/Nb}}$, and Fe substituted by Ru, $\kappa_\ell^{\text{Fe/Ru}}$. Fig. 3.16 shows the LTC of AlVFe_2 with substitutions on the different sites with different substitution concentrations. Increasing the concentration increases g_i , leading to a reduction of the LTC. It is interesting to note the importance of the choice of substitution site for AlVFe_2 . At 5 % substitution concentration the LTC is reduced by 50 % when substituting Al by Ga, 60 % when substituting V with Nb, and 65 % when substituting Fe by Ru. g_i at 5 % concentration is 0.21, 0.07, and 0.07 when substituting Al by Ga, V by Nb, and Fe by Ru, respectively. The high g_i for Al substituted by Ga comes from the large mass difference between the Al and Ga atoms when compared to the other substitutions. Although the highest g_i is for substitution on the Al-site, it results in the lowest LTC reduction. At 15 % substitution concentration $\kappa_\ell^{\text{Fe/Ru}} = 12.5 \text{ W/Km}$, which is the lowest calculated LTC. In this case, the LTC is reduced by 72 %.

When studying the effect of substitutions at 10 % for the half-Heusler class of compounds in Article II, we found that for most compounds the optimal

3. Results and Discussion

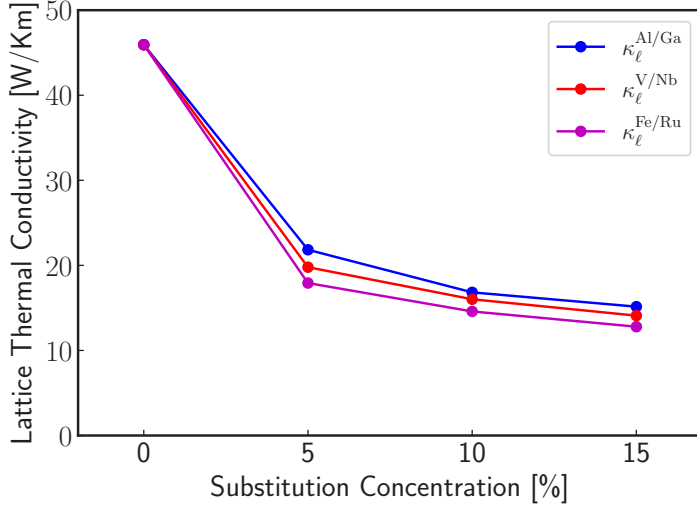


Figure 3.16: Lattice thermal conductivity at 300 K (vertical axis) for AlVFe_2 with different substitution concentrations (horizontal axis) and sites.

substitution site was the site hosting the heaviest atom. This is also the case for AlVFe_2 . Additionally, we observed that compounds with large mass differences of the atoms in the parent compound could have a large reduction in LTC with substitutions. AlVFe_2 has a relatively large mass difference with $M_{\text{Fe}}/M_{\text{Al}} \approx 2$. Although in this case study there are no other full-Heuslers to compare with, AlVFe_2 has a sizeable reduction in the LTC with substitutions, but not as large as some of the half-Heuslers. One such compound was LiBSi , which had an 80 % reduction in the LTC when substituting 10 % Si with Ge. LiBSi was also one of the half-Heuslers with the largest mass difference, $M_{\text{Si}}/M_{\text{Li}} \approx 4$.

Fig. 3.17 shows the intrinsic LTC, $\kappa_{\ell}^{\text{int}}$, and LTC with substitutions at 15 % concentration for different temperatures. $\kappa_{\ell}^{\text{int}}$ is reduced by 62 % when going from 300 K to 800 K, while $\kappa_{\ell}^{\text{Fe/Ru}}$ is reduced 49 %. This shows that substitutions are more effective at lower temperatures, $\kappa_{\ell}^{\text{Fe/Ru}}$ is 72 % lower than $\kappa_{\ell}^{\text{int}}$ at 300 K, while $\kappa_{\ell}^{\text{Fe/Ru}}$ is 63 % lower at 800 K.

With optimal site substitution (substituting Fe by Ru at 15 % concentration), the LTC is not low when compared to high-performance TE compounds (often < 1 W/Km around room temperature), and not as low as the target fixed value of 4 W/Km from Article IV. Further reduction of the LTC can be done by introducing gbs in calculations. The LTC including gb-scattering is labeled $\kappa_{\ell}^{\text{gb}}$. The $\kappa_{\ell}^{\text{gb}}$ with 300, 200, and 100 nm gbs is 37.1, 34.6, and 29.2 W/Km, respectively, at 300 K. The reduction with 100 nm gbs is 36 % when compared to $\kappa_{\ell}^{\text{int}}$. The sizes and shapes of gbs vary between compounds, and submicrometer-sized grains have been reported for AlVFe_2 [131]. Using 100 nm gbs can be

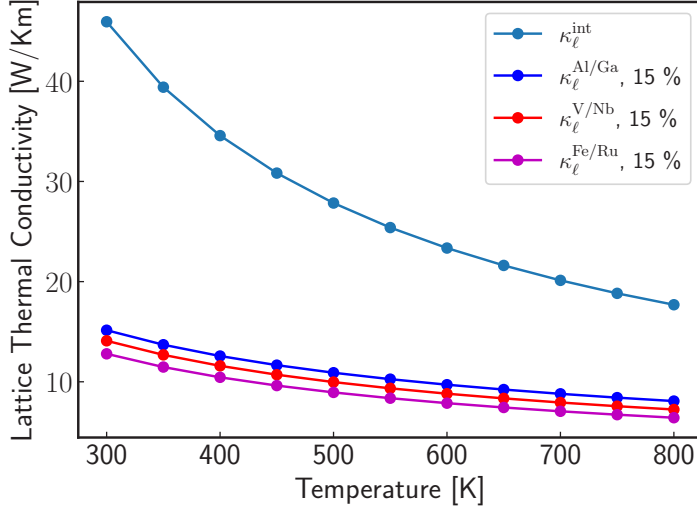


Figure 3.17: Lattice thermal conductivity for AlVFe_2 as a function of temperature. The temperature is indicated on the horizontal axis and the LTC on the vertical axis.

used as a possible best-case scenario for reducing the LTC and is used in the following.

Fig 3.18 displays the LTC with a combination of scattering from substitutions and gbs. The yellow square indicates $\kappa_{\ell}^{\text{int}}$. Substitution of Fe with 15 % Ru and 100 nm gbs leads to a 84 % reduction of the LTC. The LTC is not as low as 4 W/Km as assumed in Article IV, AlVFe_2 achieves an LTC of ~ 7 W/Km. With an electronic thermal conductivity of ~ 1 W/Km [132], the resulting thermoelectric figure of merit for AlVFe_2 becomes $ZT = 0.16$ at 300 K. Although the LTC reduction for AlVFe_2 is sizeable based on these calculations, it does not attain a high ZT .

The spectral LTC can be used to analyze the phonons that are responsible for heat transport, and also which phonons are suppressed with additional scattering mechanisms. Fig. 3.19 displays the spectral LTC for AlVFe_2 in four cases: 1) with no additional scattering mechanisms, 2) with scattering from 100 nm gbs, 3) with 15 % substitution of Fe by Ru, and 4) with scattering from 100 nm gbs and 15 % substitution. The dashed lines correspond to the cumulative LTC, which is the integral of the spectral LTC. The cumulative LTC at a given energy can be interpreted as how much LTC is accumulated from phonons with energy less than the given energy. A comparison of the DOS (Fig. 3.15) with the spectral $\kappa_{\ell}^{\text{int}}$ shows that although the phonon DOS is relatively low energies less than 20 meV, these phonons contribute significantly to the LTC. At 20 meV the cumulative $\kappa_{\ell}^{\text{int}}$ is ~ 22 W/Km. In the opposite case, phonons around 50 meV have a high DOS, while their contribution to LTC is low. The difference of the spectral $\kappa_{\ell}^{\text{int}}$ and

3. Results and Discussion

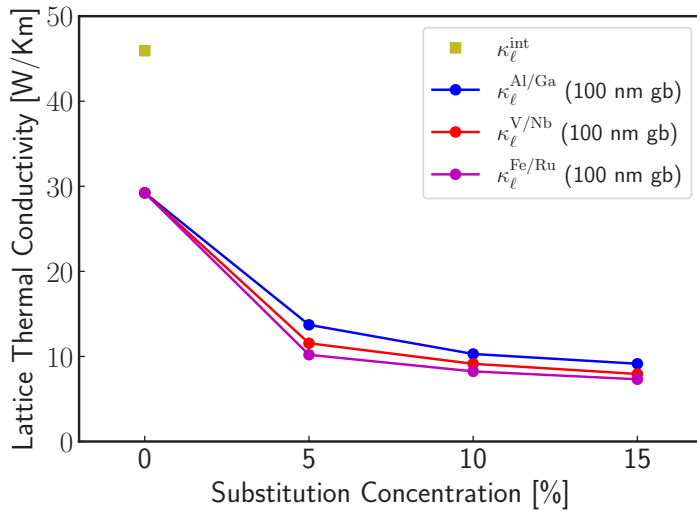


Figure 3.18: Lattice thermal conductivity at 300 K (vertical axis) with 100 nm grain boundaries for AlVFe_2 with different substitution concentrations (horizontal axis) and sites.

$\kappa_{\ell}^{\text{int}} + 100 \text{ nm gbs}$ (green and orange lines) shows that gb-scattering is effective in the range 1-10 meV, while phonons in the range 20-40 meV are less affected. The introduction of substitutions, $\kappa_{\ell}^{\text{Fe/Ru}}$, seemingly does the opposite; lower energy phonons are less affected, while higher energy phonons are suppressed. In the case of $\kappa_{\ell}^{\text{Fe/Ru}}$, almost no contribution to $\kappa_{\ell}^{\text{Fe/Ru}}$ occurs over 30 meV. Lastly, when combining the two scattering mechanisms, $\kappa_{\ell}^{\text{Fe/Ru}} + 100 \text{ nm gb}$, both low and high-energy phonons are effectively suppressed. Most of the remaining heat-conducting phonons are mid-energy phonons in the range 5-20 meV.

In conclusion, the LTC of AlVFe_2 can be reduced from ~ 46 to $\sim 7 \text{ W/Km}$ at 300 K when substituting Fe with 15 % Ru and introducing 100 nm gbs. Such a large reduction could be the difference between a mediocre and promising thermoelectric compound, although an LTC of $\sim 7 \text{ W/Km}$ is still too high for AlVFe_2 to be a promising thermoelectric compound.

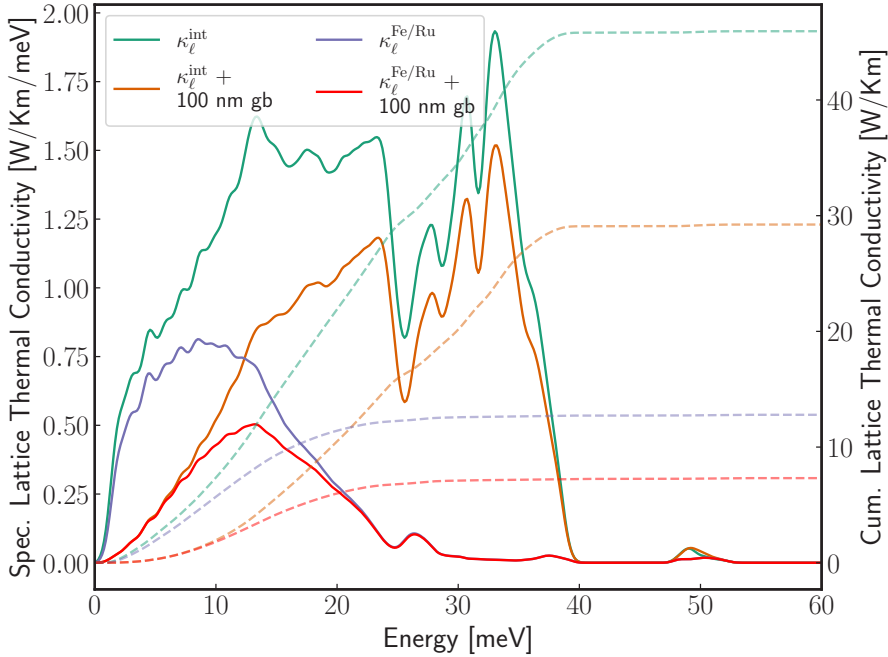


Figure 3.19: Spectral LTC (solid lines, left vertical axis) and cumulative LTC (dashed lines, right vertical axis) as a function of phonon energy for AlVFe_2 at 300 K.

3.2.3 Lattice Thermal Conductivity Calculations: Comparison with Experiment

Screening for low LTC (κ_{ℓ}) compounds is often done based on calculations as investigated in previous subsections. In the end, it is the experimental lattice thermal conductivity, $\kappa_{\ell}^{\text{exp}}$, or total thermal conductivity, κ^{exp} , that determines the real-world thermoelectric efficiency observed in applications. It is therefore interesting to assess the accuracy of theoretical intrinsic κ_{ℓ} obtained with the TDEP method, $\kappa_{\ell}^{\text{TDEP}}$, when compared to experimental values.

Fig. 3.20 shows a comparison between $\kappa_{\ell}^{\text{TDEP}}$ and $\kappa_{\ell}^{\text{exp}}$ or κ^{exp} for a set of insulators and semiconductors. The calculations include three-phonon scattering. For compounds displayed with κ^{exp} the references do not list the electrical conductivity, σ , which would have enabled estimation of the electronic thermal conductivity, κ_e , and κ_{ℓ} , using the Wiedemann-Franz law. For insulators and semiconductors it is typical that κ_{ℓ} is higher than κ_e , and in general $\kappa = \kappa_{\ell} + \kappa_e > \kappa_{\ell}$. $\kappa_{\ell}^{\text{TDEP}}$ is higher than κ^{exp} and $\kappa_{\ell}^{\text{exp}}$ except in the case of BP. For TiNiSn (Ref. [66]), ZrNiSn, NbCoSn, and VFeSb the average absolute relative error between $\kappa_{\ell}^{\text{exp}}$ and $\kappa_{\ell}^{\text{TDEP}}$ is 55 %. The ordering of compounds from low to high $\kappa_{\ell}^{\text{TDEP}}$ is close to consistent with $\kappa_{\ell}^{\text{exp}}$. For TiNiSn and ZrNiSn

3. Results and Discussion

the ordering of $\kappa_\ell^{\text{TDEP}}$ is different compared to κ_ℓ^{exp} although it is a close call. The lowest and highest κ_ℓ^{exp} included to make the average value for TiNiSn in Fig. 3.20 are 4 W/Km [66] and 7.5 W/Km [133] while $\kappa_\ell^{\text{TDEP}} = 16.6$ W/Km.

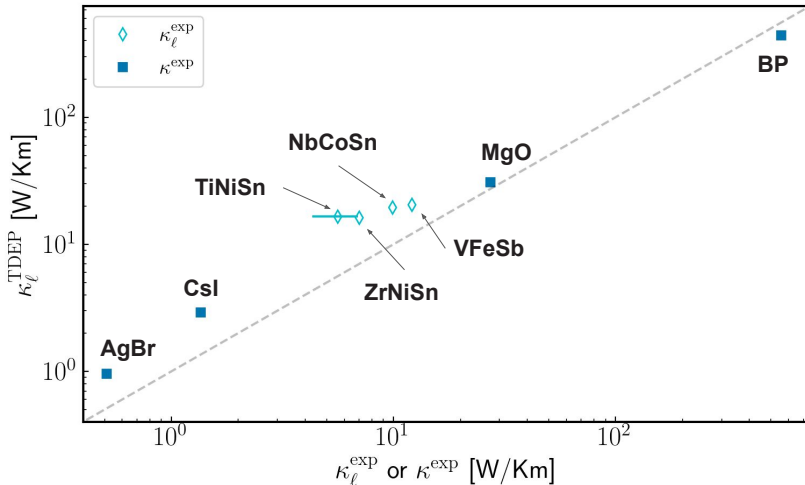


Figure 3.20: Calculated $\kappa_\ell^{\text{TDEP}}$ (this work) versus experimental κ_ℓ^{exp} or κ^{exp} . For TiNiSn [66, 133–135], ZrNiSn [66], NbCoSn [136], and VFeSb [137], κ_ℓ^{exp} is shown on the horizontal axis, while for AgBr [138], CsI [139], MgO [140], and BP [141] κ^{exp} is shown. The data is taken at 300 K except for MgO for which it is taken at 400 K. For TiNiSn an average with standard deviation is shown where data is taken at 300 K from Refs. [66, 133] and ~ 323 K from Refs. [134, 135]

There are several possible reasons for the discrepancy between the TDEP calculations and values from experiments. In these calculations, only three-phonon scattering is included. Higher-order four-phonon scattering is one mechanism that if included could reduce $\kappa_\ell^{\text{TDEP}}$, making it closer to κ_ℓ^{exp} [142]. Additional sample-dependent phonon scattering mechanisms can also partly explain the discrepancy between experimental $\kappa_\ell^{\text{TDEP}}$ and κ_ℓ^{exp} [143, 144]. Vacancies are voids in the lattice that can scatter phonons and have been observed in experimental samples such as for the half-Heuslers [145, 146]. The role of vacancies has also been studied theoretically [147], e.g. for NbCoSb [148]. This compound can have regions with Nb vacancies and impurity-rich regions with higher Nb concentrations where the Nb atoms act as interstitials. Grain boundaries are also often present in experimental samples and the size and shape of the grains impacts how much κ_ℓ is reduced [66, 149, 150].

Several other mechanisms can yield experimental results lower than calculations, such as oxygen impurities [46, 151], phase separations [152, 153], and twin boundaries [77, 78]. Incorporating the effects of additional scattering mechanisms in calculations can provide calculated $\kappa_\ell^{\text{TDEP}}$ closer to

κ_ℓ^{exp} . Including some of these mechanisms – such as four-phonon scattering – in calculations, would come at a significant computational cost [142, 147], while other mechanisms could require the development of new simulation software. Nevertheless, based on Fig. 3.20, the ranking of low to high κ_ℓ is close to consistent between calculations and experiment.

3.3 Limitations and Implications

Research projects are typically limited in terms of time and resources and also confined to certain topics. It is not uncommon to encounter challenges or problems that can not be solved within the current project. This section presents limitations of the works in this thesis. First is a discussion on the state of materials databases and screening studies. This is followed by a discussion on the limitations of lattice thermal conductivity (LTC) calculations. Afterward is a note on experimental work – in the end, experimental results prove real-world thermoelectric efficiency. Lastly, the section outlines ideas regarding machine learning (ML) screening for compounds with large LTC reduction from sublattice substitutions.

3.3.1 Databases and Screening Alternatives

The advent of materials databases with over 100 000 digital compound entries has accelerated materials discovery [154–156]. Compound entries are continuously being added to the databases but challenges remain about the "completeness" of database contents. A database materials property that becomes more reliable with more compound entries is the formation energy used to calculate the energy above the convex hull, ΔE_c . The convex hull shows what phases are most energetically stable and thermodynamic stability is achieved for $\Delta E_c = 0$ eV/atom. The construction of a reliable convex hull requires density functional theory (DFT) energies for compounds with different element concentrations as well as energies for compounds with the same element concentrations and different crystal structures. $\Delta E_c < 0.1$ eV/atom is used as a screening criterion in Article III to find compounds that are more likely to be stable. The low LTC compound Na_2TlSb has $\Delta E_c = 0$ eV/atom in Materials Project [7, 157] as of 30/6/22. There are no other compounds with the same composition in the database and two with the same elements, NaTl_2Sb and Na_6TlSb_4 . Recommending Na_2TlSb for further theoretical study or experimental realization thus poses a predicament; it could have competing phases that are more energetically favorable that are not in the database. It should be noted that the convex hull is based on 0 K calculations, and that metastable phases can exist at finite temperatures. Calculating DFT energies to populate the convex hull comes at a significant computational cost. Data sharing in material informatics thus becomes important for obtaining properties such as ΔE_c . For e.g. SiO_2 there are hundreds of compound entries in the Materials Project making a robust convex hull, providing thermodynamically stable structures.

3. Results and Discussion

Material databases contain only a fraction of all theoretically possible compounds. Populating a 4-atom structure with 50 different elements results in $50!/(50 - 4)! = 5527200$ combinations. With changes to space group and stoichiometry, the number of compounds is in the 100s of millions. Many databases have also already been screened to identify promising candidates for different applications. Jaafreh et al. tackled these challenges by constructing hypothetical prototype compounds on the form $A_mB_nX_y$ that were not available in the Inorganic Crystal Structure Database (ICSD) and used ML to predict the LTC [112], identifying new low LTC compounds. Legrain et al. [158] also conducted a study exploring compound stability of prototype structures not available in databases. They screened 71,178 half-Heusler compounds to find stable candidates. These studies highlight the benefit of compound exploration outside of conventional databases. Future studies not limited by the size and content of conventional material databases can provide hitherto unknown compounds with enticing properties.

Several choices must be made before conducting a screening study. Which compound classes should be screened, should databases be used, and if so, which databases? Different materials databases cater to different users, e.g. by containing experimental or computational results. Combining compounds from different databases is also an option. Such cases can benefit from the Open Databases Integration for Materials Design (OPTIMADE) [159, 160] which handily contains an API to connect to a variety of databases. Previous research provides valuable insights and training data for ML. Several ML studies have used LTC from the literature, and authors have shared data sets or referred to sources with calculated or experimental LTC that could serve as a starting point for developing future models [108, 110, 115, 161]. The work by Antunes et al. contains an extensive collection of sources for LTC, power factors, ZT , electronic thermal conductivity, etc. as well as sizes of the data sets [162]. Data sharing reduces the chance of duplicate calculations or experiments, and more data and data diversity could be crucial for model development. ML material informatics also benefits from code sharing and pre-trained models. Coupling pre-trained models with standardized ML packages for material science such as matminer [105] can provide an interesting path forward, where curated data, compound feature generators, and previous ML models are in one place.

This thesis is focused on the LTC, but screening based on calculated electronic properties also enables accelerated identification of thermoelectric compounds. The MaterialsProject contains ~ 35000 compounds with calculated electronic structure and band gap, $0.5 < E_g < 2.5$ eV [7]. High- ZT compounds can be explored based on their electronic structure by linking band structure characteristics and high ZT [163]. The increasing availability of electronic properties in databases promotes such screening studies. Efficient DFT calculations for electronic transport properties not available in databases also enable high-throughput screening. Compound sets in such studies could be comparable to or larger than compound sets used in DFT-based low LTC screening studies. For example, a recent report presented calculated electronic transport properties of ~ 1000 silicides [164].

There is an alternative to go straight for predicting ZT directly using ML instead of dividing the electronic and lattice transport parts [165, 166]. Jaafreh et al. [165] obtained from literature experimental ZT for 332 compounds at different temperatures, yielding 3869 samples for ML training. Using their ML model Jaafreh et al. predicted ZT for more than 10^5 hypothetical ternary compounds, $A_xB_yC_z$, $x + y + z = 1$. The model was validated using well-known thermoelectric compounds, such as half-Heusler $Zr_{0.5}Hf_{0.5}NiSn$, highlighting the potential of their method.

3.3.2 Limitations of Calculated Lattice Thermal Conductivity

LTC calculations based on DFT and the TDEP method are analyzed throughout this thesis. While Sec. 3.2.3 showed that the calculations are viable for compound ranking from low to high LTC, there is no guarantee that calculated LTC is accurate for all compounds studied.

The 0 K volume calculated with DFT is used in the LTC calculations. Neglecting the effect of thermal expansion can impact the LTC directly (Eq. 2.19), and also indirectly through the changes in force constants from differences in interatomic distances. Force constants and the equilibrium volume – and thus the LTC – can also differ notably for different DFT exchange-correlation functionals [167, 168].

The effects of four-phonon scattering on the LTC are not analyzed in this thesis. Including higher-order, four-phonon scattering mechanisms can effectively reduce LTC [142]. For the low LTC compound GeTe, the LTC was reduced from 3.8 to 1.7 W/Km – a ~ 50 % drop – when including four-phonon in addition to three-phonon scattering [169]. There is a substantial extra computational cost of calculating the fourth-order force constants required to obtain the four-phonon scattering rate. It might therefore not be suitable to include in screening studies as done in this thesis, although efforts have been made to accelerate the process of obtaining fourth- and higher-order force constants [82].

Simulating the effect of phonon scattering from sublattice substitutions through the mass-disorder scattering model implies uncertainty in the calculated LTC. The isotope-like scattering from mass-disorder is not necessarily always accurate, especially for higher concentrations [43, 44, 60]. In real-world compounds, the substituent atoms might not be evenly distributed. This is in contrast to the virtual crystal approximation and effective masses; each unit cell is identical. The accuracy of the mass-disorder model could be addressed by using the special quasirandom structure (SQS) approach [89, 170] which involves treating alloys explicitly. The assumption that it is possible to solute the substituent at the given concentration can yield erroneous results, as it could lead to phase separation in experiments [153]. The strain induced from sublattice substitutions is also not considered, and this effect has been observed significantly impact the LTC [171].

In real-world compounds, the grains vary in size and shape. Smaller grains typically scatter more phonons than large grains. The properties of the intersect at the grain boundary also affect how phonons are scattered. With the grain-

boundary scattering model in this work, the grain size is fixed, and long-wavelength phonons are scattered. Addressing the effect of grain boundaries on the LTC with more sophisticated scattering models can be useful to further understand the role of grain boundaries in reducing the LTC [172, 173], e.g. by allowing for grains of different sizes and shapes in the model.

3.3.3 Integration of Theory and Experiments

Theoretical screening studies typically involve more compounds than experimental studies. Experiments are time-consuming, expensive, and require a great deal of expertise, limiting the number of compounds studied. Screening studies often suggest compounds to be realized experimentally, underlining the importance of connecting theory and experiment. What is a promising compound with extraordinary properties if it is never realized for real-world applications? One project that combines experienced researchers from both realms is the Allotherm project [174], which is a joint computational and experimental endeavor to discover new thermoelectric compounds. This project is funded by the Research Council of Norway, an institution that invests in research and innovation on behalf of the Norwegian government. Allotherm has contributed to the attached articles, and project meetings provided valuable research input for this thesis. Feedback from experimentalists can aid theoretical compound exploration, especially when the goal is experimental realization. It can e.g. be beneficial if compounds that can not be made with current instruments or compounds that are toxic are removed from the screening study in later stages. Interdisciplinary research – where theory and experiment work in tandem – can thus be advantageous over doing an initial theoretical screening study followed by experiments.

3.3.4 Machine Learning for Lattice Thermal Conductivity with Sublattice Substitutions

Sublattice substitutions (or alloying) can be an efficient way to reduce LTC as explored in Subsection 3.2.2 and Article II. Some compounds with relatively high intrinsic LTC achieve a large reduction. These compounds might be discarded in traditional studies focusing solely on the intrinsic LTC. There have been efforts to accelerate the screening of alloyed compounds at arbitrary stoichiometry. Li et al. [175] linked the alloyed LTC to intrinsic LTC and configurational entropy. They found good agreement between predicted and experimental alloyed LTC for tetradymites and half-Heuslers. Searching for compounds that have drastic LTC reduction with alloying can also be done with ML. Such a study could be conducted similarly as in Articles I and III using an active sampling scheme. It would also be possible to train an ML model to predict the LTC with different substitution sites and concentrations. The mass-variance parameter, Eq. (2.23), and substitution concentrations can be explored as potential features in the ML model. Article II shows that the mass ratio of the atoms in the parent compound can play a role in how large the LTC reduction is. The mass ratio can therefore

be a promising feature for identifying compounds with large LTC reduction. The necessary LTC for ML training can be calculated or gathered from the literature. Calculations from Article II and the references therein can be useful as a starting point for such a study.

Chapter 4

Conclusions and Outlook

The main objective of this thesis was to identify new low lattice thermal conductivity (LTC) compounds for thermoelectric applications using computational methods. The research was based on two avenues: 1) Screening for compounds with low LTC using machine learning (ML), and 2) reducing the LTC by introducing mass-disorder and grain-boundary scattering.

4.1 Computational Screening for Promising Thermoelectric Compounds

Article I presented a case study on ML for predicting the intrinsic LTC for half-Heusler compounds using random forest regression (RFR). The lattice thermal conductivities for training were calculated based on density functional theory (DFT) and the temperature-dependent effective potential (TDEP) method. The study showed that if certain compounds were left out of the training, the R2-score was low and the model failed at identifying low LTC compounds in the test set. Active sample selection based on principal component analysis (PCA) and the features enabled the identification of crucial compounds to add to the training set. These compounds were found by evaluating compounds that were far from the training set in the space spanned by principal components. With the addition of three actively sampled compounds to the training set, the R2-score for predictions on the remaining test set drastically increased, enabling the model to correctly identify low LTC compounds.

Article III concerned an ML study with the aim to predict the LTC of cubic compounds in the Materials Project database [7]. A Gaussian Process Regression (GPR) model was trained for predicting the LTC and obtaining prediction uncertainties. The features were constructed to be compact and transferable, enabling extended screening of other compound classes or prototype structures in future studies. Active sampling based on the uncertainties indicated which compounds should be selected to be added in training. A comparison of the active sampling scheme and random sampling showed that a high R2-score could be achieved with fewer compounds in the training set when compounds were actively sampled. Thirty compounds with high uncertainty were sampled from the Materials Project and added to the training set. Two compounds with non-zero band gaps were highlighted in the study, Ca_3AsBr_3 and Na_2TlSb . Ca_3AsBr_3 was found during active sampling while Na_2TlSb was found during screening with ML predicted LTC of 0.6 W/Km. The LTC calculated with TDEP for these compounds was 0.5 and 3.8 W/Km. Na_2TlSb had low phonon group velocities in flat phonon dispersion branches as well as low phonon lifetimes. An analysis of the spectral LTC for the two compounds showed that mid-energy

4. Conclusions and Outlook

optical phonons contributed significantly to the LTC of Ca_3AsBr_3 while for Na_2TlSb almost all contribution to the LTC came from low-energy acoustic phonons.

Article IV presented a screening study of 1093 cubic compounds in the Materials Project database. The goal was to find promising thermoelectric candidates that could be missed in screening when evaluating the electronic band structure at the standard DFT level. Eight compounds had emerging electronic band gaps when assessed at the hybrid functional level of DFT. One of these compounds was AlVFe_2 , a compound with intrinsic LTC on the higher end. The potential of lowering the LTC of this compound with sublattice substitutions and grain boundaries was studied in this thesis. Ba_2HgPb had the lowest LTC among the eight compounds, although it did not display especially promising thermoelectric properties. The compounds with highest thermoelectric performance were the full-Heuslers MgSc_2Hg and Li_2CaSi , with $ZT \sim 2$.

The thesis presented a comparison of three ML models; Lasso, RFR, and support vector regression (SVR), for predicting the LTC at different temperatures. SVR performed best in cross-validation, while RFR performed best on the test set. Lasso – a linear model – performed worst in both cases. A comparison of active sampling with PCA and GPR showed that for the half-Heuslers the two schemes selected almost the same compounds; four out of the five sampled were identical. When comparing active sampling based on PCA and GPR, and random sampling, the PCA- and GPR-based sampling came out on top, providing more accurate LTC predictions for low LTC compounds.

The focus of this thesis has been on predicting the LTC directly using ML. There is an alternative to do ML for the force constants instead, potentially using similar features to those used in this work. With ML-force constants, one can obtain additional properties, such as phonon dispersions and lifetimes. These properties can be useful for interpreting lattice transport and gaining new physical insights. This thesis investigated the effect of adding different kinds of noise to the third-order force constants and then calculating the LTC. Introducing larger errors in force constants with large absolute values reduced the accuracy of the LTC. It could be necessary to differentiate strong force constants from the rest for an accurate assessment of the LTC using ML-force constants.

This work has involved predicting the LTC of cubic compounds efficiently with ML, enabling accelerated screening. While several cubic compounds have been shown to have high thermoelectric efficiency, it would be beneficial to include other types of compounds in screening, e.g. orthorhombic and layered compounds. Such compounds are readily available in materials databases. The features and active sampling scheme used in this thesis should be adaptable to different compound classes. A natural continuation of the works in this thesis would be to expand the ML method using active sampling to identify new promising low LTC candidates.

Increasing computational resources promote LTC calculations for different compound classes and for complex compounds that require additional DFT-level input. It will be beneficial to gather these calculations in a structured way in

databases to easily enable future research [162]. Increasing the size and variety of the compounds used for training can yield accurate ML models that can be employed for screening. Larger training sets can also lead to the employment of ML methods that can require large compound sets for training, such as neural networks or graph neural networks [176]. Deep neural networks have been shown to attain high accuracy in materials design and discovery, and may likely be even more prominent in the future of materials informatics [177].

4.2 Sublattice Substitutions and Grain Boundaries to Reduce Lattice Thermal Conductivity

Article II explored the potential of sublattice substitutions and grain boundaries for reducing the LTC of 122 half-Heusler compounds. The substitution concentration was set to 10 %, and grain boundaries of 50 nm and 100 nm were used. The choice of substitution site was found to be important for optimally reducing LTC – i.e. substituting on the site with the heaviest atom was optimal for 102 of the 122 compounds. Some of the half-Heuslers with medium-to-high intrinsic LTC, such as AlSiLi and TiNiPb, had a substantial LTC reduction of ~ 70 % with optimal substitution. It was shown that compounds with a large atomic mass difference in the parent compound can have especially large LTC reduction with substitutions. Introducing grain boundaries in addition to substitutions enabled four compounds with medium-to-high intrinsic LTC; LiZnSb, AlSiLi, BiPdSc, and VRhSn, to achieve LTC of < 2.2 W/Km at 300 K.

In this thesis, the LTC of full-Heusler AlVFe₂ including sublattice substitutions and grain boundaries was studied. Al was substituted by Ga, V by Nb, and Fe by Ru with concentrations of 5, 10, and 15 %. The largest LTC reduction was obtained using 15 % Ru substitutions, resulting in a 72 % drop. It was shown that the LTC can be reduced from ~ 46 W/Km to ~ 7 W/Km – an 84 % drop – with 15 % Ru-substitution and 100 nm grain boundaries. Based on the spectral LTC, the grain boundaries scattered low-energy acoustic phonons, and substitutions scattered high-energy phonons.

The thesis highlighted the importance of including additional scattering mechanisms when screening for compounds with low LTC. Discarding thermoelectric compounds with relatively high intrinsic LTC – but favorable electronic properties – can result in promising candidates being missed. Coupling screening of electronic properties together with a scheme to identify compounds with large LTC reduction with additional scattering mechanisms can be an approach to identify new high-efficiency thermoelectric candidates. With agile screening methods future studies can take advantage of prototype structures during screening, and e.g. take temperature, doping concentration, and substitution concentration as free parameters. The prospect of discovering thermoelectric materials that have high efficiency while simultaneously being cheap, non-toxic, and easy to manufacture remains bright.

Bibliography

- [1] Might, Matt. *The illustrated guide to a Ph.D.* <https://matt.might.net/articles/phd-school-in-pictures/>. Accessed: 2022-08-1.
- [2] Grimenes, Øven Andreas. *Screening of thermoelectric performance of half-Heusler materials and their alloys from first principles.* <https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/2829928>. Accessed: 2022-10-29.
- [3] Morgado, Antonio et al. “The allure of rock crystal in Copper Age southern Iberia: Technical skill and distinguished objects from Valencina de la Concepción (Seville, Spain)”. In: *Quat.* vol. 424 (2015). DOI: [10.1016/j.quaint.2015.08.004](https://doi.org/10.1016/j.quaint.2015.08.004).
- [4] Bernhard, Hans. *Wikimedia Commons.* https://commons.wikimedia.org/wiki/File:King_David_in_Augsburg_Cathedral_light.JPG. Accessed: 2022-10-28.
- [5] Pablo, Juan J. de et al. “The Materials Genome Initiative, the interplay of experiment, theory and computation”. In: *Curr. Opin. Solid State Mater. Sci.* vol. 18, no. 2 (2014), pp. 99–117. DOI: <https://doi.org/10.1016/j.cossms.2014.02.003>.
- [6] Pablo, Juan J. et al. “New frontiers for the materials genome initiative”. In: *npj Comput. Mater.* vol. 5, no. 1 (2019). DOI: [10.1038/s41524-019-0173-4](https://doi.org/10.1038/s41524-019-0173-4).
- [7] Jain, Anubhav et al. “The Materials Project: A materials genome approach to accelerating materials innovation”. In: *APL Mater.* vol. 1, no. 1 (2013), p. 011002. DOI: [10.1063/1.4812323](https://doi.org/10.1063/1.4812323).
- [8] Gutierrez Moreno, Julio et al. “A review of recent progress in thermoelectric materials through computational methods”. In: *Mater. Renew.* vol. 9, no. 3 (2020), p. 16. DOI: [10.1007/s40243-020-00175-5](https://doi.org/10.1007/s40243-020-00175-5).
- [9] Alberi, Kirstin et al. “The 2019 materials by design roadmap”. In: *J. Phys. D* vol. 52, no. 1 (2018), p. 013001. DOI: [10.1088/1361-6463/aad926](https://doi.org/10.1088/1361-6463/aad926).
- [10] Kozinsky, Boris and Singh, David J. “Thermoelectrics by Computational Design: Progress and Opportunities”. In: *Annu. Rev. Mater. Res.* vol. 51, no. 1 (2021), pp. 565–590. DOI: [10.1146/annurev-matsci-100520-015716](https://doi.org/10.1146/annurev-matsci-100520-015716).
- [11] Nozariasbmarz, Amin et al. “Review of wearable thermoelectric energy harvesting: From body temperature to electronic systems”. In: *Appl. Energy* vol. 258 (2020), p. 114069. DOI: <https://doi.org/10.1016/j.apenergy.2019.114069>.

- [12] Snyder, G. Jeffrey and Toberer, Eric S. “Complex thermoelectric materials”. In: *Nat. Mater.* vol. 7, no. 2 (2008), pp. 105–114. DOI: 10.1038/nmat2090.
- [13] Wei, Jiangtao et al. “Review of current high-ZT thermoelectric materials”. In: *J. Mater. Sci.* vol. 55, no. 5895 (2020), pp. 12642–12704. DOI: 10.1126/science.1158899.
- [14] Hasan, Md. Nazibul et al. “Inorganic thermoelectric materials: A review”. In: *International Journal of Energy Research* vol. 44, no. 8 (2020), pp. 6170–6222. DOI: <https://doi.org/10.1002/er.5313>.
- [15] Gayner, Chhatrasal and Kar, Kamal K. “Recent advances in thermoelectric materials”. In: *Prog. Mater. Sci.* vol. 83 (2016), pp. 330–382. DOI: <https://doi.org/10.1016/j.pmatsci.2016.07.002>.
- [16] *Restriction of Hazardous Substances in Electrical and Electronic Equipment (RoHS)*. https://environment.ec.europa.eu/topics/waste-and-recycling/rohs-directive_en. Accessed: 2022-04-10.
- [17] Kim, Hee Seok et al. “Relationship between thermoelectric figure of merit and energy conversion efficiency”. In: *Proc. Natl. Acad. Sci.* vol. 112, no. 27 (2015), pp. 8205–8210. DOI: 10.1073/pnas.1510231112.
- [18] Markov, Maxime et al. “Semi-metals as potential thermoelectric materials: case of HgTe”. In: *Sci. Rep.* vol. 8, no. 1 (2018). DOI: 10.1038/s41598-018-28043-3.
- [19] Peng, Bo et al. “High thermoelectric performance of Weyl semimetal TaAs”. In: *Nano Energy* vol. 30 (2016), pp. 225–234. DOI: <https://doi.org/10.1016/j.nanoen.2016.10.016>.
- [20] Zhou, Tong et al. “Enhanced thermoelectric properties of the Dirac semimetal Cd₃As₂”. In: *Inorg. Chem. Front.* vol. 3 (12 2016), pp. 1637–1643. DOI: 10.1039/C6QI00383D.
- [21] Zhou, Chongjian et al. “Polycrystalline SnSe with a thermoelectric figure of merit greater than the single crystal”. In: *Nat. Mater.* vol. 20 (2021). DOI: 10.1038/s41563-021-01064-6.
- [22] Zhu, Hangtian et al. “Discovery of TaFeSb-based half-Heuslers with high thermoelectric performance”. In: *Nat. Commun.* vol. 10, no. 1 (2019). DOI: 10.1038/s41467-018-08223-5.
- [23] Fiorentini, Vincenzo et al. “High thermoelectric figure of merit and thermopower in layered perovskite oxides”. In: *Phys. Rev. Mater.* vol. 3, no. 2 (2019). DOI: 10.1103/physrevmaterials.3.022401.
- [24] Tan, Huan et al. “Rapid preparation of Ge_{0.9}Sb_{0.1}Te_{1+x} via unique melt spinning: Hierarchical microstructure and improved thermoelectric performance”. In: *J. Alloys Compd.* vol. 774 (2019), pp. 129–136. DOI: 10.1016/j.jallcom.2018.09.144.

- [25] Zhao, Lidong et al. “Ultralow Thermal Conductivity and High Thermoelectric Figure of Merit in SnSe Crystals”. In: *Nature* vol. 508 (2014), pp. 373–7. DOI: [10.1038/nature13184](https://doi.org/10.1038/nature13184).
- [26] Xiao, Yu et al. “Origin of low thermal conductivity in SnSe”. In: *Phys. Rev. B* vol. 94 (12 2016), p. 125203. DOI: [10.1103/PhysRevB.94.125203](https://doi.org/10.1103/PhysRevB.94.125203).
- [27] Qin, Yuanhao et al. “Doping Effect on Cu₂Se Thermoelectric Performance: A Review”. In: *Materials* vol. 13 (2020). DOI: [10.3390/ma13245704](https://doi.org/10.3390/ma13245704).
- [28] Liu, Wei-Di, Yang, Lei, and Chen, Zhi-Gang. “Cu₂Se thermoelectrics: property, methodology, and device”. In: *Nano Today* vol. 35 (2020), p. 100938. DOI: <https://doi.org/10.1016/j.nantod.2020.100938>.
- [29] Sulaiman, S. et al. “Review on grain size effects on thermal conductivity in ZnO thermoelectric materials”. In: *RSC Adv.* vol. 12 (9 2022), pp. 5428–5438. DOI: [10.1039/D1RA06133J](https://doi.org/10.1039/D1RA06133J).
- [30] Sulaiman, Suraya et al. “Review of the Nanostructuring and Doping Strategies for High-Performance ZnO Thermoelectric Materials”. In: *Crystals* vol. 12, no. 8 (2022), p. 1076. DOI: [10.3390/cryst12081076](https://doi.org/10.3390/cryst12081076).
- [31] Quinn, Robert J. and Bos, Jan-Willem G. “Advances in half-Heusler alloys for thermoelectric power generation”. In: *Mater. Adv.* vol. 2 (19 2021), pp. 6246–6266. DOI: [10.1039/D1MA00707F](https://doi.org/10.1039/D1MA00707F).
- [32] Naydenov, Genadi et al. “Huge power factor in p-type half-Heusler alloys NbFeSb and TaFeSb”. In: *J. Phys. Mater.* vol. 2 (2019). DOI: [10.1088/2515-7639/ab16fb](https://doi.org/10.1088/2515-7639/ab16fb).
- [33] Hoat, D.M. “Electronic structure and thermoelectric properties of Ta-based half-Heusler compounds with 18 valence electrons”. In: *Comput. Mater. Sci.* vol. 159 (2019), pp. 470–477. DOI: <https://doi.org/10.1016/j.commatsci.2018.12.039>.
- [34] Yuan, H.M. et al. “Promising thermoelectric performance of full-Heusler compound Sr₂AuBi”. In: *Phys. Lett. A* vol. 404 (2021), p. 127413. DOI: <https://doi.org/10.1016/j.physleta.2021.127413>.
- [35] Anand, Shashwat et al. “Thermoelectric transport of semiconductor full-Heusler VFe₂Al”. In: *J. Mater. Chem. C* vol. 8 (2020), pp. 10174–10184. DOI: [10.1039/D0TC02659J](https://doi.org/10.1039/D0TC02659J).
- [36] Haque, Md Azimul et al. “Halide Perovskites: Thermal Transport and Prospects for Thermoelectricity”. In: *Adv. Sci.* vol. 7, no. 10 (2020), p. 1903389. DOI: <https://doi.org/10.1002/adv.201903389>.
- [37] Zhou, Yingzhi et al. “Recent progress of halide perovskites for thermoelectric application”. In: *Nano Energy* vol. 94 (2022), p. 106949. DOI: <https://doi.org/10.1016/j.nanoen.2022.106949>.
- [38] Huang, Lihong et al. “Recent progress in half-Heusler thermoelectric materials”. In: *Mater. Res. Bull.* vol. 76 (2016), pp. 107–112. DOI: <https://doi.org/10.1016/j.materresbull.2015.11.032>.

- [39] Gaultois, Michael W. and Sparks, Taylor D. “How much improvement in thermoelectric performance can come from reducing thermal conductivity?” In: *Appl. Phys. Lett.* vol. 104, no. 11 (2014), p. 113906. DOI: 10.1063/1.4869232.
- [40] Srivastava, G.P. *The Physics of Phonons*. Taylor & Francis, 1990.
- [41] *Speeds of sound of the elements*. https://en.wikipedia.org/wiki/Speeds_of_sound_of_the_elements. Accessed: 2022-12-1.
- [42] Hopcroft, Matthew A., Nix, William D., and Kenny, Thomas W. “What is the Young’s Modulus of Silicon?” In: *J. Microelectromech. Syst.* vol. 19, no. 2 (2010), pp. 229–238. DOI: 10.1109/JMEMS.2009.2039697.
- [43] Tamura, Shin-ichiro. “Isotope scattering of dispersive phonons in Ge”. In: *Phys. Rev. B* vol. 27 (1983), pp. 858–866. DOI: 10.1103/PhysRevB.27.858.
- [44] Tamura, Shin-ichiro. “Isotope scattering of large-wave-vector phonons in GaAs and InSb: Deformation-dipole and overlap-shell models”. In: *Phys. Rev. B* vol. 30 (1984), pp. 849–854. DOI: 10.1103/PhysRevB.30.849.
- [45] Eliassen, Simen N. H. et al. “Lattice thermal conductivity of $\text{Ti}_x\text{Zr}_y\text{Hf}_{1-x-y}\text{NiSn}$ half-Heusler alloys calculated from first principles: Key role of nature of phonon modes”. In: *Phys. Rev. B* vol. 95 (2017), p. 045202. DOI: 10.1103/PhysRevB.95.045202.
- [46] Slack, G.A. “Nonmetallic crystals with high thermal conductivity”. In: *J. Phys. Chem. Solids* vol. 34, no. 2 (1973), pp. 321–335. DOI: [https://doi.org/10.1016/0022-3697\(73\)90092-9](https://doi.org/10.1016/0022-3697(73)90092-9).
- [47] Cahill, David G., Watson, S. K., and Pohl, R. O. “Lower limit to the thermal conductivity of disordered crystals”. In: *Phys. Rev. B* vol. 46 (1992), pp. 6131–6140. DOI: 10.1103/PhysRevB.46.6131.
- [48] Anderson, Orson L. “A simplified method for calculating the debye temperature from elastic constants”. In: *J. Phys. Chem. Solids* vol. 24, no. 7 (1963), pp. 909–917. DOI: [https://doi.org/10.1016/0022-3697\(63\)90067-2](https://doi.org/10.1016/0022-3697(63)90067-2).
- [49] Jia, Tiantian, Chen, Gang, and Zhang, Yongsheng. “Lattice thermal conductivity evaluated using elastic properties”. In: *Phys. Rev. B* vol. 95 (2017), p. 155206. DOI: 10.1103/PhysRevB.95.155206.
- [50] Delaire, Olivier et al. “Giant anharmonic phonon scattering in PbTe.” In: *Nat. Mater.* vol. 10 8 (2011), pp. 614–9. DOI: 10.1038/nmat3035.
- [51] Dangić, Đorđe et al. “The origin of the lattice thermal conductivity enhancement at the ferroelectric phase transition in GeTe”. In: *npj Comput. Materi.* vol. 7 (2021). DOI: 10.1038/s41524-021-00523-7.
- [52] Pal, Koushik et al. “Accelerated discovery of a large family of quaternary chalcogenides with very low lattice thermal conductivity”. In: *npj Comput. Mater* vol. 7, 82 (2021), p. 82. DOI: 10.1038/s41524-021-00549-x.

- [53] Sales, B.C. et al. “Atomic Displacement Parameters and the Lattice Thermal Conductivity of Clathrate-like Thermoelectric Compounds”. In: *J. Solid State Chem.* vol. 146, no. 2 (1999), pp. 528–532. DOI: <https://doi.org/10.1006/jssc.1999.8354>.
- [54] Tadano, Terumasa, Gohda, Yoshihiro, and Tsuneyuki, Shinji. “Impact of Rattlers on Thermal Conductivity of a Thermoelectric Clathrate: A First-Principles Study”. In: *Phys. Rev. Lett.* vol. 114, no. 9 (2015), p. 095501. DOI: [10.1103/PhysRevLett.114.095501](https://doi.org/10.1103/PhysRevLett.114.095501).
- [55] Sales, B. C., Mandrus, D., and Williams, R. K. “Filled Skutterudite Antimonides: A New Class of Thermoelectric Materials”. In: *Science* vol. 272, no. 5266 (1996), pp. 1325–1328. DOI: [10.1126/science.272.5266.1325](https://doi.org/10.1126/science.272.5266.1325).
- [56] Feng, Zhenzhen et al. “Characterization of rattling in relation to thermal conductivity: ordered half-Heusler semiconductors”. In: *Phys. Rev. B* vol. 101, no. 6 (2020), p. 064301. DOI: [10.1103/PhysRevB.101.064301](https://doi.org/10.1103/PhysRevB.101.064301).
- [57] Dong, Jinfeng et al. “Reducing Lattice Thermal Conductivity of MnTe by Se Alloying toward High Thermoelectric Performance”. In: *ACS Appl. Mater. Interfaces.* vol. 11, no. 31 (2019). PMID: 31305979, pp. 28221–28227. DOI: [10.1021/acsami.9b10207](https://doi.org/10.1021/acsami.9b10207).
- [58] Heinrich, Christophe et al. “Effect of isovalent substitution on the thermoelectric properties of the $\text{Cu}_2\text{ZnGeSe}_{4-x}\text{S}_x$ series of solid solutions”. In: *J. Am. Chem. Soc.* vol. 136 (2013). DOI: [10.1021/ja410753k](https://doi.org/10.1021/ja410753k).
- [59] Sun, Hui, Lu, Xu, and Morelli, Donald T. “Isovalent substitutes play in different ways: Effects of isovalent substitution on the thermoelectric properties of $\text{CoSi}_{0.98}\text{B}_{0.02}$ ”. In: *J. Appl. Phys.* vol. 120, no. 3 (2016), p. 035107. DOI: [10.1063/1.4959209](https://doi.org/10.1063/1.4959209).
- [60] Tranås, Rasmus, Løvvik, Ole Martin, and Berland, Kristian. “Attaining Low Lattice Thermal Conductivity in Half-Heusler Sublattice Solid Solutions: Which Substitution Site Is Most Effective?” In: *Electron. Mater.* vol. 3, no. 1 (2022), pp. 1–14. DOI: [10.3390/electronicmat3010001](https://doi.org/10.3390/electronicmat3010001).
- [61] Li, Chao et al. “High Thermoelectric Performance Achieved in Sb-Doped GeTe by Manipulating Carrier Concentration and Nanoscale Twin Grains”. In: *Materials* vol. 15, no. 2 (2022). DOI: [10.3390/ma15020406](https://doi.org/10.3390/ma15020406).
- [62] Namiki, Hiromasa et al. “Enhancement and manipulation of the thermoelectric properties of n-type argyrodite Ag_8SnSe_6 with ultralow thermal conductivity by controlling the carrier concentration through Ta doping”. In: *AIP Adv.* vol. 11, no. 7 (2021), p. 075125. DOI: [10.1063/5.0056533](https://doi.org/10.1063/5.0056533).
- [63] Liang, Xin. “Impact of grain boundary characteristics on lattice thermal conductivity: A kinetic theory study on ZnO”. In: *Phys. Rev. B* vol. 95 (15 2017), p. 155313. DOI: [10.1103/PhysRevB.95.155313](https://doi.org/10.1103/PhysRevB.95.155313).

- [64] Zhao, Huaizhou et al. “Engineering the Thermoelectric Transport in Half-Heusler Materials through a Bottom-Up Nanostructure Synthesis”. In: *Adv. Energy Mater.* vol. 7 (2017). DOI: [10.1002/aenm.201700446](https://doi.org/10.1002/aenm.201700446).
- [65] Cho, Junsang et al. “Ti Addition Effect on the Grain Structure Evolution and Thermoelectric Transport Properties of $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{NiSn}_{0.98}\text{Sb}_{0.02}$ Half-Heusler Alloy”. In: *Materials* vol. 14 (2021), p. 4029. DOI: [10.3390/ma14144029](https://doi.org/10.3390/ma14144029).
- [66] Schrade, Matthias et al. “The role of grain boundary scattering in reducing the thermal conductivity of polycrystalline XNiSn ($\text{X} = \text{Hf}, \text{Zr}, \text{Ti}$) half-Heusler alloys”. In: *Sci. Rep.* vol. 7 (2017). DOI: [10.1038/s41598-017-14013-8](https://doi.org/10.1038/s41598-017-14013-8).
- [67] Osei-Agyemang, Eric, Adu, Challen Enniful, and Balasubramanian, Ganesh. “Ultralow lattice thermal conductivity of chalcogenide perovskite CaZrSe_3 contributes to high thermoelectric figure of merit”. In: *npj Comput. Mater.* vol. 5, no. 1 (2019). DOI: [10.1038/s41524-019-0253-5](https://doi.org/10.1038/s41524-019-0253-5).
- [68] Zhang, Ping et al. “Reduced lattice thermal conductivity of perovskite-type high-entropy ($\text{Ca}_{0.25}\text{Sr}_{0.25}\text{Ba}_{0.25}\text{RE}_{0.25}$)/ TiO_3 ceramics by phonon engineering for thermoelectric applications”. In: *J. Alloys Compd.* vol. 898 (2022), p. 162858. DOI: <https://doi.org/10.1016/j.jallcom.2021.162858>.
- [69] He, Huifang et al. “Role of vacancy defects on the lattice thermal conductivity in In_2O_3 thermoelectric nanocrystals: a positron annihilation study”. In: *J. Mater. Sci.* vol. 53 (2018). DOI: [10.1007/s10853-018-2544-5](https://doi.org/10.1007/s10853-018-2544-5).
- [70] Kim, Chang-eun et al. “Effect of Vacancy Distribution on the Thermal Conductivity of Ga_2Te_3 and Ga_2Se_3 ”. In: *J. Electron. Mater.* vol. 40 (2011). DOI: [10.1007/s11664-010-1479-7](https://doi.org/10.1007/s11664-010-1479-7).
- [71] Duan, Bo et al. “Regulation of oxygen vacancy and reduction of lattice thermal conductivity in ZnO ceramic by high temperature and high pressure method”. In: *Ceram. Int.* vol. 46, no. 16, Part A (2020), pp. 26176–26181. DOI: <https://doi.org/10.1016/j.ceramint.2020.07.115>.
- [72] Fu, Chenguang et al. “Realizing high figure of merit in heavy-band p -type half-Heusler thermoelectric materials”. In: *Nat. Commun.* vol. 6 (2015), p. 8144. DOI: [10.1038/ncomms9144](https://doi.org/10.1038/ncomms9144).
- [73] Fan, D. D. et al. “A first-principles study of the effects of electron–phonon coupling on the thermoelectric properties: a case study of the SiGe compound”. In: *J. Mater. Chem. A* vol. 6, no. 25 (2018), pp. 12125–12131. DOI: [10.1039/C8TA01806E](https://doi.org/10.1039/C8TA01806E).
- [74] Karthikeyan, Vaithinathan et al. “Dislocation-induced ultra-low lattice thermal conductivity in rare earth doped $\beta - \text{Zn}_4\text{Sb}_3$ ”. In: *Scr. Mater.* vol. 174 (2020), pp. 95–101. DOI: <https://doi.org/10.1016/j.scriptamat.2019.08.037>.

- [75] Amouyal, Yaron. “A Practical Approach to Evaluate Lattice Thermal Conductivity in Two-Phase Thermoelectric Alloys for Energy Applications”. In: *Materials* vol. 10, no. 4 (2017), p. 386. DOI: [10.3390/ma10040386](https://doi.org/10.3390/ma10040386).
- [76] Babaei, Hasan, McGaughey, Alan J. H., and Wilmer, Christopher E. “Effect of pore size and shape on the thermal conductivity of metal-organic frameworks”. In: *Chem. Sci.* vol. 8, no. 1 (2017), pp. 583–589. DOI: [10.1039/c6sc03704f](https://doi.org/10.1039/c6sc03704f).
- [77] Mao, Jun et al. “Phonon scattering by nanoscale twin boundaries”. In: *Nano Energy* vol. 32 (2017), pp. 174–179. DOI: <https://doi.org/10.1016/j.nanoen.2016.12.026>.
- [78] Bagri, Akbar et al. “Thermal transport across Twin Grain Boundaries in Polycrystalline Graphene from Nonequilibrium Molecular Dynamics Simulations”. In: *Nano Lett.* vol. 11, no. 9 (2011), pp. 3917–3921. DOI: [10.1021/nl202118d](https://doi.org/10.1021/nl202118d).
- [79] Togo, A and Tanaka, I. “First principles phonon calculations in materials science”. In: *Scr. Mater.* vol. 108 (2015), pp. 1–5. DOI: [10.1016/j.scriptamat.2015.07.021](https://doi.org/10.1016/j.scriptamat.2015.07.021).
- [80] Togo, Atsushi, Chaput, Laurent, and Tanaka, Isao. “Distributions of phonon lifetimes in Brillouin zones”. In: *Phys. Rev. B* vol. 91 (9 2015), p. 094306. DOI: [10.1103/PhysRevB.91.094306](https://doi.org/10.1103/PhysRevB.91.094306).
- [81] Li, Wu et al. “ShengBTE: a solver of the Boltzmann transport equation for phonons”. In: *Comp. Phys. Commun.* vol. 185 (2014), pp. 1747–1758. DOI: [10.1016/j.cpc.2014.02.015](https://doi.org/10.1016/j.cpc.2014.02.015).
- [82] Eriksson, Fredrik, Fransson, Erik, and Erhart, Paul. “The Hiphive Package for the Extraction of High-Order Force Constants by Machine Learning”. In: *Adv. Theory Simul.* vol. 2, no. 5 (2019), p. 1800184. DOI: [10.1002/adts.201800184](https://doi.org/10.1002/adts.201800184).
- [83] Hellman, O., Abrikosov, I. A., and Simak, S. I. “Lattice dynamics of anharmonic solids from first principles”. In: *Phys. Rev. B* vol. 84, no. 18 (2011), p. 180301. DOI: [10.1103/PhysRevB.84.180301](https://doi.org/10.1103/PhysRevB.84.180301).
- [84] Hellman, Olle and Broido, David A. “Phonon thermal transport in Bi_2Te_3 from first principles”. In: *Phys. Rev. B* vol. 90, no. 13 (2014), p. 134309. DOI: [10.1103/PhysRevB.90.134309](https://doi.org/10.1103/PhysRevB.90.134309).
- [85] Hellman, Olle and Abrikosov, I. A. “Temperature-dependent effective third-order interatomic force constants from first principles”. In: *Phys. Rev. B* vol. 88, no. 14 (2013), p. 144301. DOI: [10.1103/PhysRevB.88.144301](https://doi.org/10.1103/PhysRevB.88.144301).
- [86] Dewandre, Antoine et al. “Two-Step Phase Transition in SnSe and the Origins of its High Power Factor from First Principles”. In: *Phys. Rev. Lett.* vol. 117 (27 2016), p. 276601. DOI: [10.1103/physrevlett.117.276601](https://doi.org/10.1103/physrevlett.117.276601).
- [87] Shulumba, Nina, Hellman, Olle, and Minnich, Austin J. “Intrinsic localized mode and low thermal conductivity of PbSe ”. In: *Phys. Rev. B* vol. 95 (1 2017), p. 014302. DOI: [10.1103/PhysRevB.95.014302](https://doi.org/10.1103/PhysRevB.95.014302).

- [88] Shulumba, Nina, Hellman, Olle, and Minnich, Austin J. “Lattice Thermal Conductivity of Polyethylene Molecular Crystals from First-Principles Including Nuclear Quantum Effects.” In: *Phys. Rev. Lett.* vol. 119 18 (2017), p. 185901. DOI: [10.1103/physrevlett.119.185901](https://doi.org/10.1103/physrevlett.119.185901).
- [89] Shulumba, Nina. *Vibrations in Solids: From first principles lattice dynamics to high temperature phase stability*. Linköping University Electronic Press, 2015, p. 94. DOI: [10.3384/diss.diva-122949](https://doi.org/10.3384/diss.diva-122949).
- [90] *Temperature dependent effective potential 1.1*. <https://ollehellman.github.io/index.html>. Accessed: 2022-21-10.
- [91] West, D. and Estreicher, S. K. “First-Principles Calculations of Vibrational Lifetimes and Decay Channels: Hydrogen-Related Modes in Si”. In: *Phys. Rev. Lett.* vol. 96 (11 2006), p. 115504. DOI: [10.1103/PhysRevLett.96.115504](https://doi.org/10.1103/PhysRevLett.96.115504).
- [92] West, D. and Estreicher, S. K. “First-Principles Calculations of Vibrational Lifetimes and Decay Channels: Hydrogen-Related Modes in Si”. In: *Phys. Rev. Lett.* vol. 96 (11 2006), p. 115504. DOI: [10.1103/PhysRevLett.96.115504](https://doi.org/10.1103/PhysRevLett.96.115504).
- [93] Kresse, G. and Hafner, J. “*Ab initio* molecular dynamics for liquid metals”. In: *Phys. Rev. B* vol. 47, no. 1 (1993), pp. 558–561. DOI: [10.1103/PhysRevB.47.558](https://doi.org/10.1103/PhysRevB.47.558).
- [94] Kresse, G. and Furthmüller, J. “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set”. In: *Comput. Mater. Sci.* vol. 6, no. 1 (1996), pp. 15–50. DOI: [10.1016/0927-0256\(96\)00008-0](https://doi.org/10.1016/0927-0256(96)00008-0).
- [95] Kresse, G. and Furthmüller, J. “Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set”. In: *Phys. Rev. B* vol. 54, no. 16 (1996), pp. 11169–11186. DOI: [10.1103/PhysRevB.54.11169](https://doi.org/10.1103/PhysRevB.54.11169).
- [96] Ramprasad, Rampi et al. “Machine Learning and Materials Informatics: Recent Applications and Prospects”. In: *npj Comput. Mater.* vol. 3 (2017). DOI: [10.1038/s41524-017-0056-5](https://doi.org/10.1038/s41524-017-0056-5).
- [97] Rajan, Krishna. “Materials Informatics: The Materials “Gene” and Big Data”. In: *Annu. Rev. Mater. Res.* vol. 45, no. 1 (2015), pp. 153–169. DOI: [10.1146/annurev-matsci-070214-021132](https://doi.org/10.1146/annurev-matsci-070214-021132).
- [98] Ward, Logan and Wolverton, Chris. “Atomistic calculations and materials informatics: A review”. In: *Curr. Opin. Solid State Mater. Sci.* vol. 21, no. 3 (2017), pp. 167–176. DOI: <https://doi.org/10.1016/j.cossms.2016.07.002>.
- [99] Miyazaki, Hidetoshi et al. “Machine learning based prediction of lattice thermal conductivity for half-Heusler compounds using atomic information”. In: *Sci. Rep.* (2021). DOI: [10.1038/s41598-021-92030-4](https://doi.org/10.1038/s41598-021-92030-4).

- [100] *What is Materials Informatics?* <https://citrine.io/what-is-materials-informatics-blog/>. Accessed: 2022-03-10.
- [101] Wang, Anthony Yu-Tung et al. “Machine Learning for Materials Scientists: An Introductory Guide toward Best Practices”. In: *Chem. Mater.* vol. 32, no. 12 (2020), pp. 4954–4965. DOI: 10.1021/acs.chemmater.0c01907.
- [102] Seko, Atsuto, Togo, Atsushi, and Tanaka, Isao. “Descriptors for Machine Learning of Materials Data”. In: *Nanoinformatics*. Ed. by Tanaka, Isao. Singapore: Springer Singapore, 2018, pp. 3–23. DOI: 10.1007/978-981-10-7617-6_1.
- [103] Kanda, Yosuke, Fujii, Hitoshi, and Oguchi, Tamio. “Sparse modeling of chemical bonding in binary compounds”. In: *Sci. Technol. Adv. Mater.* vol. 20, no. 1 (2019). PMID: 32082439, pp. 1178–1188. DOI: 10.1080/14686996.2019.1697858.
- [104] *mendeleev – A Python resource for properties of chemical elements, ions and isotopes, ver. 0.10.0*. <https://github.com/lmmentel/mendeleev>. 2014–.
- [105] Ward, Logan et al. “Matminer: An open source toolkit for materials data mining”. In: *Comput. Mater. Sci.* vol. 152 (2018), pp. 60–69. DOI: <https://doi.org/10.1016/j.commatsci.2018.05.018>.
- [106] Voronoi, Georges. “Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les paralléloèdres primitifs.” In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* vol. 1908, no. 134 (1908), pp. 198–287. DOI: doi:10.1515/crll.1908.134.198.
- [107] Ward, Logan et al. “Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations”. In: *Phys. Rev. B* vol. 96 (2 2017), p. 024104. DOI: 10.1103/PhysRevB.96.024104.
- [108] Chen, Lihua et al. “Machine learning models for the lattice thermal conductivity prediction of inorganic materials”. In: *Comput. Mater. Sci.* vol. 170 (2019), p. 109155. DOI: <https://doi.org/10.1016/j.commatsci.2019.109155>.
- [109] Juneja, Rinkle and Singh, Abhishek K. “Guided patchwork kriging to develop highly transferable thermal conductivity prediction models”. In: *J. Phys. Mater.* vol. 3, no. 2 (2020), p. 024006. DOI: 10.1088/2515-7639/ab78f2.
- [110] Zhu, Taishan et al. “Charting lattice thermal conductivity for inorganic crystals and discovering rare earth chalcogenides for thermoelectrics”. In: *Energy Environ. Sci.* vol. 14 (6 2021), pp. 3559–3566. DOI: 10.1039/D1EE00442E.
- [111] Carrete, Jesús et al. “Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling”. In: *Phys. Rev. X* vol. 4 (2014). DOI: 10.1103/PhysRevX.4.011019.

- [112] Jaafreh, Russlan, Kang, Yoo Seong, and Hamad, Kotiba. “Lattice Thermal Conductivity: An Accelerated Discovery Guided by Machine Learning”. In: *ACS Appl. Mater. Interfaces* vol. 13, no. 48 (2021), pp. 57204–57213. DOI: [10.1021/acsami.1c17378](https://doi.org/10.1021/acsami.1c17378).
- [113] Tranås, Rasmus et al. “Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis”. In: *Comput. Mater. Sci.* vol. 202 (2022), p. 110938. DOI: <https://doi.org/10.1016/j.commatsci.2021.110938>.
- [114] Seko, Atsuto et al. “Prediction of Low-Thermal-Conductivity Compounds with First-Principles Anharmonic Lattice-Dynamics Calculations and Bayesian Optimization”. In: *Phys. Rev. Lett.* vol. 115 (20 2015), p. 205901. DOI: [10.1103/PhysRevLett.115.205901](https://doi.org/10.1103/PhysRevLett.115.205901).
- [115] Ju, Shenghong et al. “Exploring diamond-like lattice thermal conductivity crystals via feature-based transfer learning”. In: *Phys. Rev. Mater.* vol. 5 (5 2021), p. 053801. DOI: [10.1103/PhysRevMaterials.5.053801](https://doi.org/10.1103/PhysRevMaterials.5.053801).
- [116] Knoop, Florian et al. “Anharmonicity measure for materials”. In: *Phys. Rev. Mater.* vol. 4 (8 2020), p. 083809. DOI: [10.1103/PhysRevMaterials.4.083809](https://doi.org/10.1103/PhysRevMaterials.4.083809).
- [117] Roekeghem, Ambroise van et al. “High-Throughput Computation of Thermal Conductivity of High-Temperature Solid Phases: The Case of Oxide and Fluoride Perovskites”. In: *Phys. Rev. X* vol. 6 (4 2016), p. 041061. DOI: [10.1103/PhysRevX.6.041061](https://doi.org/10.1103/PhysRevX.6.041061).
- [118] Jinnouchi, Ryosuke, Karsai, Ferenc, and Kresse, Georg. “On-the-fly machine learning force field generation: Application to melting points”. In: *Phys. Rev. B* vol. 100 (1 2019), p. 014105. DOI: [10.1103/PhysRevB.100.014105](https://doi.org/10.1103/PhysRevB.100.014105).
- [119] Schwalbe-Koda, Daniel, Tan, Aik, and Gómez-Bombarelli, Rafael. “Differentiable sampling of molecular geometries with uncertainty-based adversarial attacks”. In: *Nat. Commun.* vol. 12, no. 1 (2021). DOI: [10.1038/s41467-021-25342-8](https://doi.org/10.1038/s41467-021-25342-8).
- [120] Kleiven, David et al. “Training sets based on uncertainty estimates in the cluster-expansion method”. In: *J. Phys. Energy* vol. 3, no. 3 (2021), p. 034012. DOI: [10.1088/2515-7655/abf9ef](https://doi.org/10.1088/2515-7655/abf9ef).
- [121] Lookman, Turab et al. “Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design”. In: *npj Comput. Mater.* vol. 5 (2019). DOI: [10.1038/s41524-019-0153-8](https://doi.org/10.1038/s41524-019-0153-8).
- [122] Ebden, Mark. *Gaussian Processes: A Quick Introduction*. 2015. DOI: [10.48550/ARXIV.1505.02965](https://doi.org/10.48550/ARXIV.1505.02965).
- [123] Zhou, Fei et al. “Lattice Anharmonicity and Thermal Conductivity from Compressive Sensing of First-Principles Calculations”. In: *Phys. Rev. Lett.* vol. 113 (18 2014), p. 185501. DOI: [10.1103/PhysRevLett.113.185501](https://doi.org/10.1103/PhysRevLett.113.185501).

- [124] Fransson, Erik, Eriksson, Fredrik, and Erhart, Paul. “Efficient construction of linear models in materials modeling and applications to force constant expansions”. In: *npj Comput. Mater.* vol. 6, no. 1 (). DOI: 10.1038/s41524-020-00404-5.
- [125] Nishino, Y., Deguchi, S., and Mizutani, U. “Thermal and transport properties of the Heusler-type $\text{Fe}_2\text{VAl}_{1-x}\text{Ge}_x$ ($0 \leq x \leq 0.20$) alloys: Effect of doping on lattice thermal conductivity, electrical resistivity, and Seebeck coefficient”. In: *Phys. Rev. B* vol. 74 (11 2006), p. 115115. DOI: 10.1103/PhysRevB.74.115115.
- [126] Renard, Krystel et al. “Thermoelectric properties of the Heusler-type $\text{Fe}_2\text{VTa}_x\text{Al}_{1-x}$ alloys”. In: *J. Appl. Phys.* vol. 115, no. 3 (2014), p. 033707. DOI: 10.1063/1.4861419.
- [127] Garmroudi, F. et al. “Boosting the thermoelectric performance of Fe_2VAl -type Heusler compounds by band engineering”. In: *Phys. Rev. B* vol. 103 (8 2021), p. 085202. DOI: 10.1103/PhysRevB.103.085202.
- [128] Hinterleitner, B. et al. “Thermoelectric performance of a metastable thin-film Heusler alloy”. In: *Nature* vol. 576 (2019). DOI: 10.1038/s41586-019-1751-9.
- [129] Kundu, Ashis et al. “Ultrahigh Thermal Conductivity of θ -Phase Tantalum Nitride”. In: *Phys. Rev. Lett.* vol. 126 (11 2021), p. 115901. DOI: 10.1103/PhysRevLett.126.115901.
- [130] Peng, Wanyue et al. “An Unlikely Route to Low Lattice Thermal Conductivity: Small Atoms in a Simple Layered Structure”. In: *Joule* vol. 2, no. 9 (2018), pp. 1879–1893. DOI: <https://doi.org/10.1016/j.joule.2018.06.014>.
- [131] Mikami, M. et al. “Thermoelectric properties of tungsten-substituted Heusler Fe_2VAl alloy”. In: *J. Appl. Phys.* vol. 111, no. 9 (2012), p. 093710. DOI: 10.1063/1.4710990.
- [132] Berland, Kristian, Løvvik, Ole Martin, and Tranås, Rasmus. “Discarded gems: Thermoelectric performance of materials with band gap emerging at the hybrid-functional level”. In: *Appl. Phys. Lett.* vol. 119, no. 8 (2021), p. 081902. DOI: 10.1063/5.0058685.
- [133] Bhattacharya, S. et al. “Grain structure effects on the lattice thermal conductivity of Ti-based half-Heusler alloys”. In: *Appl. Phys. Lett.* vol. 81, no. 1 (2002), pp. 43–45. DOI: 10.1063/1.1488698.
- [134] Barczak, Sonia A. et al. “Grain-by-Grain Compositional Variations and Interstitial Metals—A New Route toward Achieving High Performance in Half-Heusler Thermoelectrics”. In: *ACS Appl. Mater. Interfaces* vol. 10, no. 5 (2018), pp. 4786–4793. DOI: 10.1021/acsami.7b14525.

- [135] Bhardwaj, A. and Misra, D. K. “Improving the thermoelectric performance of TiNiSn half-Heusler via incorporating submicron lamellae eutectic phase of Ti_{70.5}Fe_{20.5}: a new strategy for enhancing the power factor and reducing the thermal conductivity”. In: *J. Mater. Chem. A* vol. 2, no. 48 (2014), pp. 20980–20989. DOI: [10.1039/c4ta04661g](https://doi.org/10.1039/c4ta04661g).
- [136] He, Ran et al. “Enhanced thermoelectric properties of n-type NbCoSn half-Heusler by improving phase purity”. In: *APL Mater.* vol. 4, no. 10 (2016), p. 104804. DOI: [10.1063/1.4952994](https://doi.org/10.1063/1.4952994).
- [137] Fu, Chenguang et al. “Enhanced phonon scattering by mass and strain field fluctuations in Nb substituted FeVSb half-Heusler thermoelectric materials”. In: *J. Appl. Phys.* vol. 112, no. 12 (2012), p. 124915. DOI: [10.1063/1.4772605](https://doi.org/10.1063/1.4772605).
- [138] Kamran, Kashif, Anis-ur-Rehman, M, and Maqsood, Asghari. “Thermal and electrical properties of crystalline silver bromide”. In: *J. Phys. D: Appl. Phys.* vol. 40, no. 3 (2007), pp. 869–873. DOI: [10.1088/0022-3727/40/3/027](https://doi.org/10.1088/0022-3727/40/3/027).
- [139] Gerlich, D and Andersson, P. “Temperature and pressure effects on the thermal conductivity and heat capacity of CsCl, CsBr and CsI”. In: *J. Solid State Phys.* vol. 15, no. 25 (1982), pp. 5211–5222. DOI: [10.1088/0022-3719/15/25/013](https://doi.org/10.1088/0022-3719/15/25/013).
- [140] Slifka, Andrew, Filla, Bernard, and Phelps, J. “Thermal Conductivity of Magnesium Oxide From Absolute, Steady-State Measurements”. In: *J. Res. Natl. Inst. Stand. Technol.* vol. 103 (1998). DOI: [10.6028/jres.103.021](https://doi.org/10.6028/jres.103.021).
- [141] Kang, Joon Sang, Wu, Huan, and Hu, Yongjie. “Thermal Properties and Phonon Spectral Characterization of Synthetic Boron Phosphide for High Thermal Conductivity Applications”. In: *Nano Lett.* vol. 17, no. 12 (2017). PMID: 29115845, pp. 7507–7514. DOI: [10.1021/acs.nanolett.7b03437](https://doi.org/10.1021/acs.nanolett.7b03437).
- [142] Feng, Tianli, Lindsay, Lucas, and Ruan, Xiulin. “Four-phonon scattering significantly reduces intrinsic thermal conductivity of solids”. In: *Phys. Rev. B* vol. 96 (16 2017), p. 161201. DOI: [10.1103/PhysRevB.96.161201](https://doi.org/10.1103/PhysRevB.96.161201).
- [143] Xie, Hanhui et al. “The intrinsic disorder related alloy scattering in ZrNiSn half-Heusler thermoelectric materials”. In: *Sci. Rep.* vol. 4 (2014), p. 6888. DOI: [10.1038/srep06888](https://doi.org/10.1038/srep06888).
- [144] Ferluccio, Daniella A et al. “Thermal properties of TiNiSn and VFeSb half-Heusler thermoelectrics from synchrotron x-ray powder diffraction”. In: *JPhys. Energy* vol. 3, no. 3 (2021), p. 035001. DOI: [10.1088/2515-7655/abf41a](https://doi.org/10.1088/2515-7655/abf41a).
- [145] Roth, Nikolaj, Zhu, Tiejun, and Iversen, Bo B. “A simple model for vacancy order and disorder in defective half-Heusler systems”. In: *IUCrJ* vol. 7, no. 4 (2020), pp. 673–680. DOI: [10.1107/S2052252520005977](https://doi.org/10.1107/S2052252520005977).

- [146] Hazama, Hirofumi et al. “Improvement of thermoelectric properties for half-Heusler TiNiSn by interstitial Ni defects”. In: *J. Appl. Phys.* vol. 110 (2011). DOI: [10.1063/1.3633518](https://doi.org/10.1063/1.3633518).
- [147] Katre, Ankita, Carrete, Jesús, and Mingo, Natalio. “Unraveling the dominant phonon scattering mechanism in thermoelectric compound ZrNiSn”. In: *J. Mater. Chem. A* vol. 4 (2016). DOI: [10.1039/C6TA05868J](https://doi.org/10.1039/C6TA05868J).
- [148] Xia, Kaiyang et al. “Enhanced Thermoelectric Performance in 18-Electron Nb_{0.8}CoSb Half-Heusler Compound with Intrinsic Nb Vacancies”. In: *Adv. Funct. Mater.* vol. 28, no. 9 (2018), p. 1705845. DOI: <https://doi.org/10.1002/adfm.201705845>.
- [149] Luo, Ting et al. “Dopant-segregation to grain boundaries controls electrical conductivity of n-type NbCo(Pt)Sn half-Heusler alloy mediating thermoelectric performance”. In: *Acta Mater.* vol. 217 (2021), p. 117147. DOI: <https://doi.org/10.1016/j.actamat.2021.117147>.
- [150] Smith, David S. et al. “Grain boundary thermal resistance and finite grain size effects for heat conduction through porous polycrystalline alumina”. In: *Int. J. Heat Mass Transf.* vol. 121 (2018), pp. 1273–1280. DOI: <https://doi.org/10.1016/j.ijheatmasstransfer.2018.01.082>.
- [151] Zhang, Zhirui et al. “The quantitative investigation of the lattice oxygen and grain edge oxygen on the thermal conductivity of aluminum nitride ceramics”. In: *J. Eur. Ceram. Soc.* vol. 43, no. 2 (2023), pp. 313–320. DOI: <https://doi.org/10.1016/j.jeurceramsoc.2022.10.023>.
- [152] Gelbstein, Yaniv et al. “Significant lattice thermal conductivity reduction following phase separation of the highly efficient Ge_xPb_{1-x}Te thermoelectric alloys”. In: *Phys. Status Solidi B* vol. 251, no. 7 (2014), pp. 1431–1437. DOI: [10.1002/pssb.201451088](https://doi.org/10.1002/pssb.201451088).
- [153] Berche, Alexandre, Tédénac, J.-C, and Jund, Philippe. “Phase separation in the half-Heusler thermoelectric materials (Hf,Ti,Zr)NiSn”. In: *Scr. Mater.* vol. 139 (2017), pp. 122–125. DOI: [10.1016/j.scriptamat.2017.06.036](https://doi.org/10.1016/j.scriptamat.2017.06.036).
- [154] Ju, Shenghong and Shiomi, Junichiro. “Materials Informatics for Heat Transfer: Recent Progresses and Perspectives”. In: *Nanoscale Microscale Thermophys. Eng.* vol. 23, no. 2 (2019), pp. 157–172. DOI: [10.1080/15567265.2019.1576816](https://doi.org/10.1080/15567265.2019.1576816).
- [155] Choudhary, Kamal, Garrity, Kevin F, and Tavazza, Francesca. “Data-driven discovery of 3D and 2D thermoelectric materials”. In: *J. Phys. Condens. Matter* vol. 32, no. 47 (2020), p. 475501. DOI: [10.1088/1361-648x/aba06b](https://doi.org/10.1088/1361-648x/aba06b).
- [156] Zhang, Hang et al. “Machine Learning for Novel Thermal-Materials Discovery: Early Successes, Opportunities, and Challenges”. In: *ES energy environ.* vol. 2 (2018), pp. 1–8. DOI: [10.30919/esee8c209](https://doi.org/10.30919/esee8c209).

- [157] Munro, Jason et al. “An improved symmetry-based approach to reciprocal space path selection in band structure calculations”. In: *npj Comput. Mater.* vol. 6 (2020), p. 112. DOI: [10.1038/s41524-020-00383-7](https://doi.org/10.1038/s41524-020-00383-7).
- [158] Legrain, Fleur et al. “Materials Screening for the Discovery of New Half-Heuslers: Machine Learning versus ab Initio Methods”. In: *J. Phys. Chem. B* vol. 122, no. 2 (2018). PMID: 28742351, pp. 625–632. DOI: [10.1021/acs.jpcc.7b05296](https://doi.org/10.1021/acs.jpcc.7b05296).
- [159] Andersen, Casper W. et al. “OPTIMADE, an API for exchanging materials data”. In: *Sci.* vol. 8, no. 1 (2021). DOI: [10.1038/s41597-021-00974-z](https://doi.org/10.1038/s41597-021-00974-z).
- [160] Andersen, Casper et al. *The OPTIMADE Specification*. en. 2021. DOI: [10.5281/ZENODO.4195050](https://doi.org/10.5281/ZENODO.4195050).
- [161] Gaultois, Michael W. et al. “A recommendation engine for suggesting unexpected thermoelectric chemistries”. In: *APL Mater.* vol. 4, no. 5 (2016), p. 053213. DOI: [10.1063/1.4952607](https://doi.org/10.1063/1.4952607).
- [162] Antunes, Luis M. et al. “Machine Learning Approaches for Accelerating the Discovery of Thermoelectric Materials”. In: *Machine Learning in Materials Informatics: Methods and Applications*. Chap. 1, pp. 1–32. DOI: [10.1021/bk-2022-1416.ch001](https://doi.org/10.1021/bk-2022-1416.ch001).
- [163] Park, Junsoo et al. “Optimal band structure for thermoelectrics with realistic scattering and bands”. In: *npj Comput. Mater.* vol. 7, no. 1 (2021). DOI: [10.1038/s41524-021-00512-w](https://doi.org/10.1038/s41524-021-00512-w).
- [164] Løvvik, Ole Martin, Flage-Larsen, Espen, and Skomedal, Gunstein. “Screening of thermoelectric silicides with atomistic transport calculations”. In: *J. Appl. Phys.* vol. 128, no. 12 (2020), p. 125105. DOI: [10.1063/5.0008198](https://doi.org/10.1063/5.0008198).
- [165] Jaafreh, Russlan et al. “A deep learning perspective into the figure-of-merit of thermoelectric materials”. In: *Mater. Lett.* vol. 319 (2022), p. 132299. DOI: <https://doi.org/10.1016/j.matlet.2022.132299>.
- [166] Na, Gyoung S. and Chang, Hyunju. “A public database of thermoelectric materials and system-identified material representation for data-driven discovery”. In: *npj Comput. Mater.* vol. 8, no. 1 (2022). DOI: [10.1038/s41524-022-00897-2](https://doi.org/10.1038/s41524-022-00897-2).
- [167] Jain, Ankit and McGaughey, Alan J.H. “Effect of exchange–correlation on first-principles-driven lattice thermal conductivity predictions of crystalline silicon”. In: *Comput. Mater. Sci.* vol. 110 (2015), pp. 115–120. DOI: <https://doi.org/10.1016/j.commatsci.2015.08.014>.
- [168] Skelton, Jonathan M. et al. “Influence of the exchange-correlation functional on the quasi-harmonic lattice dynamics of II–VI semiconductors”. In: *J. Chem. Phys.* vol. 143, no. 6 (2015), p. 064710. DOI: [10.1063/1.4928058](https://doi.org/10.1063/1.4928058).
- [169] Xia, Yi and Chan, Maria K. Y. “Anharmonic stabilization and lattice heat transport in rocksalt -GeTe”. In: *Appl. Phys. Lett.* vol. 113, no. 19 (2018), p. 193902. DOI: [10.1063/1.5048814](https://doi.org/10.1063/1.5048814).

- [170] Røe, Ingeborg Treu. *Finite temperature ab initio simulations of the lattice thermal conductivity of XNiBi, X = (Sc, Y, La), based half-Heuslers*. <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2453064>. Accessed: 2022-10-29.
- [171] Arrighoni, Marco et al. “First-principles quantitative prediction of the lattice thermal conductivity in random semiconductor alloys: The role of force-constant disorder”. In: *Phys. Rev. B* vol. 98 (11 2018), p. 115205. DOI: 10.1103/PhysRevB.98.115205.
- [172] Kazan, M. and Volz, S. “Calculation of the lattice thermal conductivity in granular crystals”. In: *J. Appl. Phys.* vol. 115, no. 7 (2014), p. 073509. DOI: 10.1063/1.4866362.
- [173] Sood, Aditya et al. “Direct Visualization of Thermal Conductivity Suppression Due to Enhanced Phonon Scattering Near Individual Grain Boundaries”. In: *Nano Lett.* vol. 18, no. 6 (2018). PMID: 29631399, pp. 3466–3472. DOI: 10.1021/acs.nanolett.8b00534.
- [174] *Allotherm - High-throughput alloy design of superior thermoelectric materials*. <https://www.sintef.no/prosjekter/2021/allotherm-high-throughput-alloy-design-of-superior-thermoelectric-materials/>. Accessed: 2022-06-07.
- [175] Li, Mengke et al. “Predicting the lattice thermal conductivity of alloyed compounds from the perspective of configurational entropy”. In: *npj Comput. Mater.* vol. 8, 75 (2022), p. 75. DOI: 10.1038/s41524-022-00771-1.
- [176] Reiser, Patrick et al. “Graph neural networks for materials science and chemistry”. In: *Commun. Mater.* vol. 3, no. 1 (2022). DOI: 10.1038/s43246-022-00315-6.
- [177] Agrawal, Ankit and Choudhary, Alok. “Deep materials informatics: Applications of deep learning in materials science”. In: *MRS Commun.* vol. 9, no. 3 (2019), pp. 779–792. DOI: 10.1557/mrc.2019.73.

Papers

Paper I

Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis

R. Tranås, O. M. Løvvik, O. Tomic, and K. Berland

PUBLISHED. *Comput. Mater. Sci.*, 202, 110938, (2022)



Editor's choice



Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis

Rasmus Tranås^{a,*}, Ole Martin Løvvik^{b,c}, Oliver Tomic^d, Kristian Berland^a

^a Department of Mechanical Engineering and Technology Management, Norwegian University of Life Sciences, Drabakveien 31, Ås, 1432, Ås, Norway

^b Centre for Materials Science and Nanotechnology, Department of Physics, University of Oslo, Sem Selanders vei 26, Oslo, 0371, Oslo, Norway

^c Materials Physics, SINTEF, Forskningsveien 1, Oslo, 0373, Oslo, Norway

^d Department of Data Science, Norwegian University of Life Sciences, Drabakveien 31, Ås, 1432, Ås, Norway

ARTICLE INFO

Keywords:

Lattice thermal conductivity
Density functional theory
Machine learning
Principal component analysis
Half-Heusler compounds

ABSTRACT

Low lattice thermal conductivity is essential for high thermoelectric performance of a material. Lattice thermal conductivity is often computed using density functional theory (DFT), typically at a high computational cost. Training machine learning models to predict lattice thermal conductivity could offer an effective procedure to identify low lattice thermal conductivity compounds. However, in doing so, we must face the fact that such compounds can be quite rare and distinct from those in a typical training set. This distinctness can be problematic as standard machine learning methods are inaccurate when predicting properties of compounds with features differing significantly from those in the training set. By computing the lattice thermal conductivity of 122 half-Heusler compounds, using the temperature-dependent effective potential method, we generate a data set to explore this issue. We first show how random forest regression can fail to identify low lattice thermal conductivity compounds with random selection of training data. Next, we show how active selection of training data using feature and principal component analysis can be used to improve model performance and the ability to identify low lattice thermal conductivity compounds. Lastly, we find that active learning without the use of DFT-based features can be viable as a quicker way of selecting samples.

1. Introduction

With their ability to convert heat to electricity, thermoelectrics find use in several niche technologies ranging from wine coolers, hiking stoves with mobile phone chargers, and radioisotope thermoelectric (TE) generators used to power e.g. the *Curiosity* Mars rover. Thermoelectrics could also contribute to reducing global greenhouse gas emissions through waste heat recovery, but their role is currently limited by the modest efficacy realized in devices [1,2]. Another limitation is the fact that several state-of-the-art TE materials contain toxic or rare elements [3,4]. Finding new TE materials has therefore gathered much scientific interest in recent years [5].

The efficiency of TE materials is conventionally given by the dimensionless figure of merit, which is expressed as $ZT = \sigma S^2 T / (\kappa_e + \kappa_l)$, where σ is the electrical conductivity, S is the Seebeck coefficient, T is the absolute temperature, κ_e is the electronic thermal conductivity, and κ_l is the lattice thermal conductivity. High ZT requires both a high power factor, $\mathcal{P} = \sigma S^2$, and low total thermal conductivity. In non-metals, κ_l is typically much larger than κ_e , but in heavily doped

semiconductors, κ_l and κ_e can be more comparable in size [6,7]; nonetheless, a low κ_l is still typically needed for achieving high ZT .

High-throughput screening based on first-principle calculations have in recent years been much used in the search for new TE materials [8–14]. Many studies focus on electronic properties and use simple models or estimates of κ_l . One reason for this is that computing κ_l comes at a significant computational cost. The cost arises because accounting for the phonon–phonon interactions due to the anharmonicity of the lattice vibrations requires obtaining third-order force constants extracted from a large number of supercell-based density functional theory (DFT) calculations [15–17]. For this reason, machine learning (ML) methods are increasingly supplementing first-principles based calculations for predicting κ_l [18–25]. Pre-trained ML models can in turn also be made available in convenient web-based applications [26].

The half-Heusler (HH) compounds are a class of cubic compounds with three atoms in the primitive cell, belonging to the $F43m$ space-group. As shown in Fig. 1, the XZ sublattice forms a rocksalt structure,

* Corresponding author.

E-mail address: rasmus.andre.tranas@nmbu.no (R. Tranås).

<https://doi.org/10.1016/j.commatsci.2021.110938>

Received 8 July 2021; Received in revised form 22 September 2021; Accepted 28 September 2021
0927-0256/© 2021 Elsevier B.V. All rights reserved.

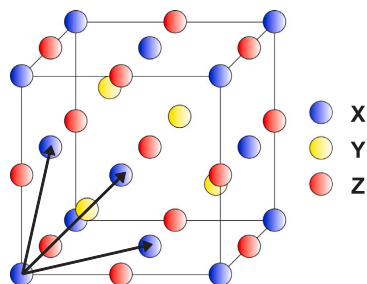


Fig. 1. The HH crystal structure, displayed as the unit cell. The primitive cell is made from the three atoms X , Y , and Z .

while the YZ sublattice forms a zinc-blende structure [27–29]. Several HH compounds have a high power factor in combination with relatively low κ_F , making HHs attractive for TE applications [8,13,30–36]. Recently, Feng et al. [37] used DFT-based calculations to show that the four HH compounds: CdPNa, BaBiK, LaRhTe, and LaPtSb have very low κ_F . LaPtSb and BaBiK also have promising electronic transport properties in addition to low κ_F and could have a ZT competitive with top performing TE materials [38–40]. Because κ_F of these four compounds is much lower than for typical HHs, the set of HH compounds presents itself as a dataset well suited for investigating ML methods to separate low and high κ_F compounds. The high symmetry of HHs also reduces the computational cost of calculating κ_F compared to more complex systems such as layered compounds and compounds with distorted symmetries [41–43]. The reduced cost allows us to generate both training and test sets for assessing κ_F .

Our study is based on 122 HH compounds for which κ_F is computed explicitly using DFT. The compounds are based on a combination of dynamically stable HHs, 54 from groups 4-9-15 (Ti,Zr,Hf)(Co,Rh,Ir)(As,Sb,Bi), 4-10-14 (Ti,Zr,Hf)(Ni,Pd,Pt)(Ge,Sn,Pb), and 48 HHs from groups 5-8-15 (V,Nb,Ta)(Fe,Ru,Os)(As,Sb,Bi) and 5-9-14 (V,Nb,Ta)(Co,Rh,Ir)(Ge,Sn,Pb). The last 20 HHs are the remaining stable compounds studied by Feng et al. [37], based on a revision of the 75 stable HHs identified by Carrete et al. [18].

2. Methods

2.1. Lattice thermal conductivity

DFT calculations in this work are done with the VASP [44–46] software package using the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation for solids, PBEsol [47,48]. The plane-wave energy cutoff is set to 600 eV. For relaxations, we use an $11 \times 11 \times 11$ k-point sampling of the Brillouin zone. The electronic self-consistent loop is iterated until the energy difference falls below 10^{-6} eV, while ionic positions are relaxed until forces fall below $1 \text{ meV}/\text{\AA}$. The lattice thermal conductivity, κ_F , is calculated with the temperature-dependent effective potential (TDEP) method [16,49], taking into account three-phonon and isotope-phonon scattering [50,51]. Fifty configurations based on $3 \times 3 \times 3$ repetitions of the primitive cell are used to obtain second- and third-order force constants. The atomic configurations are taken from a fixed-temperature canonical ensemble at 300 K, where the zero-point motion of the phonons is matched with the Debye temperature [52]. The Debye temperature is obtained from the Voigt approximation of the bulk and shear moduli [53]. A $3 \times 3 \times 3$ k-point grid is used for the supercell DFT force calculations. We employ a cutoff for second-order pair-interactions of 7 \AA while for third-order pair-interactions, the cutoff is set slightly larger than half the width of the supercell (i.e. 6.1 \AA for NbCoGe). For the calculation of κ_F , the reciprocal space is discretized on a $35 \times 35 \times 35$ q-point grid. In a convergence study for NbCoGe, we find these cutoffs to give a numerical error of κ_F less than 3%.

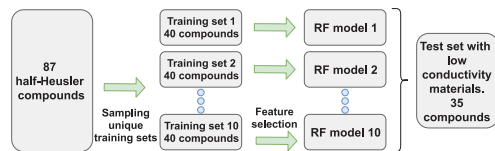


Fig. 2. Flowchart for building the baseline model.

2.2. Machine learning model

Random forest (RF) regression is a non-linear ML method used in industry and academia alike [54]. An ensemble of decision trees forms the RF model, where each tree is trained on a subset of randomly chosen features and training samples. This randomness makes RF less prone to overfitting. RF has been shown to perform well in earlier ML studies involving the lattice thermal conductivity [23]. In the RF regression, a given input sample is sorted in each of the decision trees based on its features, so that in a given tree, the sample is assigned to a $\kappa_F^{(i,j)}$ in the training set. Finally, the predicted outcome is given by the mean $\langle \kappa_F^{(i,j)} \rangle$ of the predictions of the ensemble of decision trees.

In ML, failing to identify key features can result in overfitting and reduce method interpretability [55,56]. Feature selection is here performed using exhaustive feature selection (EFS) in combination with RF regression. EFS assesses the predictive performance of every subset of extracted features and finds the features that give the best outcome of a chosen performance metric. We here choose to use Spearman rank correlation as the metric with five-fold cross-validation, as this correlation measures the predicted ranking of compounds. This brute-force approach carries a significant computational cost, but with the limited number of features in our study, this cost is small compared to that of computing κ_F . EFS is done with the MLXTEND [57] code, while RF regression is done using SCIKIT-LEARN [58]. In the RF model, for each set of features hyperparameters are optimized using a hyperparameter grid search.

Fig. 2 shows a flowchart for the baseline model. In the model, 87 of the 122 compounds are semi-randomly selected as the training pool for ML, while 35 are left out to provide a test set for model assessment. By semi-randomly, we refer to the fact the five lowest κ_F compounds are in the test set. We make this choice to emulate a not too improbable scenario that could easily arise for larger material classes when only modest-size training sets are used. From the training pool, 10 unique training sets of 40 compounds are selected randomly. The models are retrained based on the features obtained with EFS for final model evaluation. Our baseline model predictions are given by the average of the predictions of these 10 RF models.

In the active sampling scheme, we use principal component analysis (PCA) with the HOGGORM [59] package to identify compounds possessing combinations of feature values that are distinct from those in the training pool. PCA accounts for correlations between features by constructing orthogonal principal components (PC) as linear combinations of feature vectors in feature space. The PCs are oriented in the direction of maximum variance and the features are centered and scaled to unit variance. The PCA analysis is based on all compounds in the study. Using PCA, we identify three compounds, BaBiK, CdPNa, and LaPtSb, that are needed to cover the feature space mapped out by the first two PCs. These three are subsequently included in the training sets from the baseline model, such that the 10 training sets for the active sampling model contain 43 compounds that are used with RF and EFS.

Our study is based on 14 features: 9 are tabulated while 5 are obtained from low-cost DFT calculations. Two of these, the volume of the relaxed primitive unit cell, V , and corresponding mass density, ρ , could in principle have been obtained from typical tabulated data, standard experiment, or in the absence of such data, from ML models [60–62].

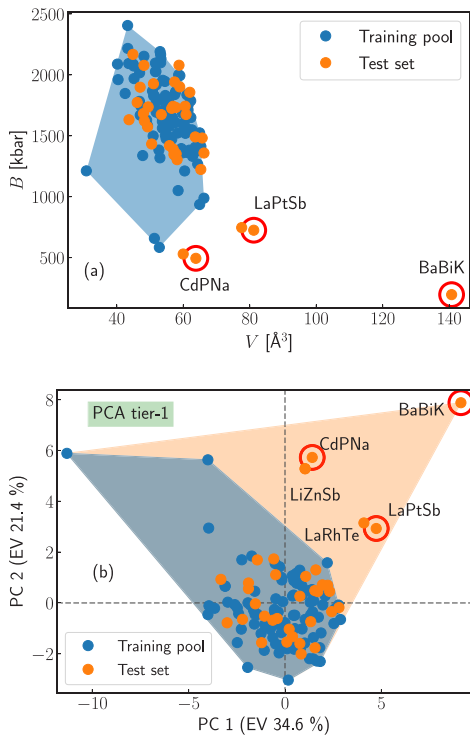


Fig. 4. (a) Scatter plot for the compounds, with V on the horizontal axis and B on the vertical axis. (b) PCA plot for the compounds using tier-1 features. The first PC is shown on the horizontal axis and the second is on the vertical axis. The blue convex hull is the area of the PC space spanned by the 87 materials in the training pool, while the orange convex hull is the area spanned with inclusion of the test compounds.

4. Results and discussion: Machine learning

4.1. Using principal component analysis for diversifying training sets

Fig. 4(a) shows the position of the compounds in the space spanned by V and B , two features which are known to correlate with the lattice thermal conductivity [18,20,71]. In general, low V materials tend to have less compressed acoustical phonon band structures, increasing the phonon group velocity [72] and thus also the lattice thermal conductivity. Higher B can also be related to stiffer atomic bonds and increased phonon velocities. This is reflected in a Spearman correlation of -0.69 between $\kappa_{\ell}^{\text{TDEP}}$ and V and 0.58 between $\kappa_{\ell}^{\text{TDEP}}$ and B for the compounds in the training pool. The plot shows that some of the compounds in the test set fall outside the convex hull spanned by V and B of the compounds in the training pool. Including such outliers in the training sets could result in more accurate ML predictions. However, as B and V also have a Spearman correlation of -0.51 , —i.e. higher V tend to relate to less stiff bonds and thus lower B — solely relying on these two features could risk missing important compounds and correlations. This motivates the use of PCA, which offers a more systematic procedure to take all features and their correlations into consideration.

Fig. 4(b) indicates the position of the compounds in the test set and training pool in the subspace spanned by the two first PCs. This subspace accounts for 56.0% of the cumulative explained variance (EV) of the feature space. Mapping these two features back to the original feature space, we find the cumulative explained variances of V and B

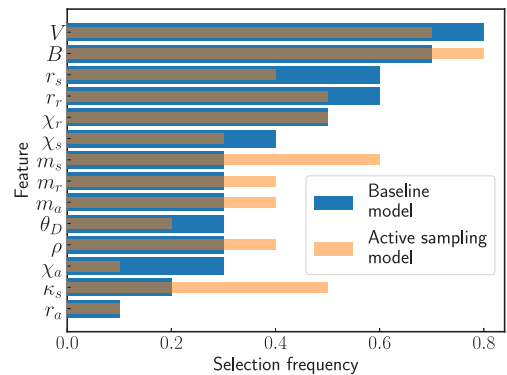


Fig. 5. Selection frequencies for the 14 features in the EFS using RF regression for the baseline (blue bars) and active sampling (orange bars) models.

to be 78.6% and 77.2%. A comparison of the two convex hulls shows that the low κ_{ℓ} compounds lie outside of convex hull spanned by the training pool.

While PCA can support a human-guided selection of training sets, we choose to formalize this in a systematic procedure that is better suited for automation of the active sampling. The specific compounds to be included are determined iteratively by identifying the compound in the test set with the largest Euclidean distance in the PC space to the closest compound in the current training pool until a marked drop in distance arises. The procedure identifies that three additional compounds, BaBiK, CdPNa, and LaPtSb, should be included in the training process. These compounds are highlighted with red circles in Fig. 4, and Fig. 7 shows the distance in PC space after each iteration.

4.2. Exhaustive feature selection analysis

Fig. 5 compares the EFS feature selection frequency of the baseline and active sampling ML models. The baseline and active sampling models use on average 5.7 and 5.9 features out of the 14 potential features, respectively. The relatively few features selected is in line with the recent results of Miyazaki et al. [60] finding that using a limited subset of features gives the best ML performance, which can be linked to the fact that redundant features can cause overfitting. In both the baseline and active sampling models, B and V are the most frequently selected features, in agreement with their high Spearman correlation with $\kappa_{\ell}^{\text{TDEP}}$.

There are some notable differences between the EFS for the active sampling model and the baseline model. In particular, the selection frequency of m_s increases from 0.3 to 0.6 for the active sampling model. This result reflects that the variation of masses in the primitive cell is linked to low lattice thermal conductivity, such as for BaBiK. The selection frequencies for κ_s and B also increase for the active sampling model.

4.3. Enhanced machine learning performance with active sampling

Fig. 6(a) compares the predictions of the baseline model and the active sampling model on a logarithmic scale as used in the training. The error bars indicate the standard deviation of the predictions of the 10 models. The figure shows that the active sampling model has a superior ability to identify the compounds with low $\kappa_{\ell}^{\text{TDEP}}$. Predictions for the three compounds found with PCA, highlighted with red circles, are not provided for the active sampling model as they are included in the training sets of the model. Fig. 6(b) shows the corresponding comparison with a linear scale, with compounds sorted according to

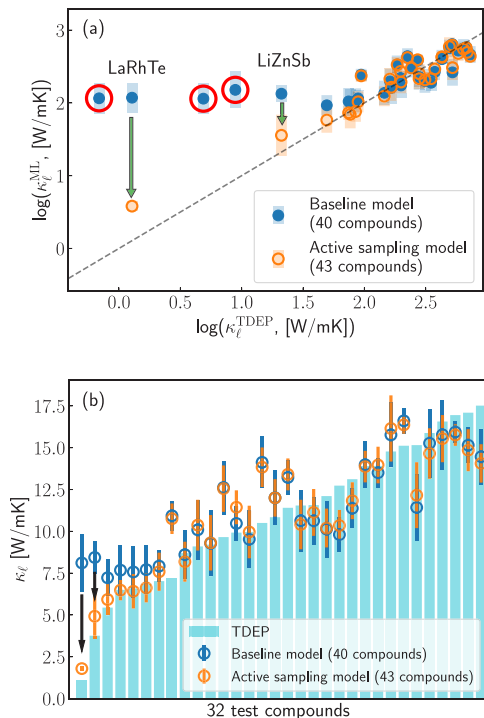


Fig. 6. (a) Parity plot for predictions made on the test set. The horizontal axis shows $\log(\kappa_{\ell}^{\text{TDEP}})$ at 500 K, while the vertical axis shows the predictions. Blue (orange) circles indicate the predictions made by the baseline (active sampling) model. (b) Corresponding comparison for $\kappa_{\ell}^{\text{TDEP}}$. The turquoise bars indicate $\kappa_{\ell}^{\text{TDEP}}$ at 500 K.

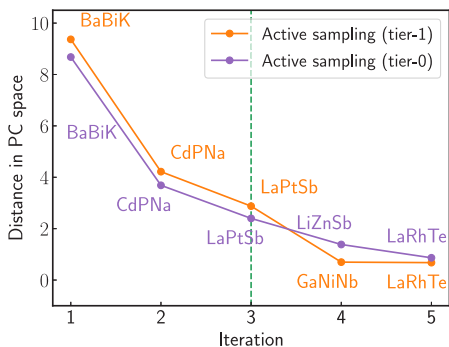


Fig. 7. Distance in PC space between the compound that is farthest from the training pool and its closest training pool neighbor after each iteration of sampling. Orange indicates PC space distance in the space spanned with tier-1 features, and purple shows the results for tier-0 features. The green line indicates the point at which we stop the inclusion of more compounds.

$\kappa_{\ell}^{\text{TDEP}}$. In most of the cases, the active sampling model predictions, $\kappa_{\ell}^{\text{AS}}$, are higher than $\kappa_{\ell}^{\text{TDEP}}$ for low $\kappa_{\ell}^{\text{TDEP}}$ compounds, and vice versa for high $\kappa_{\ell}^{\text{TDEP}}$ compounds. This is also seen in the logarithmic scale of Fig. 6(a). Even if the numerical precision of the active sampling model for the compounds with low $\kappa_{\ell}^{\text{TDEP}}$ is quite modest, which can be linked to the limited sampling in this region of feature space, the model identifies

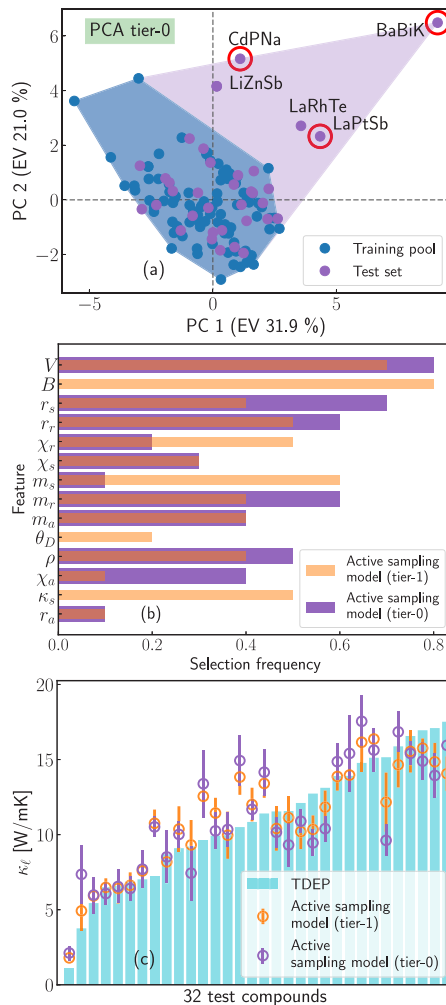


Fig. 8. (a) PCA plot for the compounds using tier-0 features. The horizontal axis shows the first PC while the vertical axis shows the second. (b) Selection frequencies from the EFS using RF regression for the active sampling models using tier-0 and tier-1 features. (c) Predicted and TDEP lattice thermal conductivity using active sampling models based on tier-0 and tier-1 features.

Table 1

Performance metrics for predicting $\log(\kappa_{\ell}^{\text{TDEP}})$ for the 32 test compounds using the active sampling and baseline models. The metrics are: R2-score, root-mean-square error (RMSE), Spearman correlation, and Pearson correlation. The standard deviations are in parenthesis.

	Active	Baseline
R2	0.84 (0.03)	0.36 (0.13)
RMSE	0.21 (0.02)	0.43 (0.04)
Spearman	0.85 (0.04)	0.79 (0.07)
Pearson	0.93 (0.02)	0.64 (0.09)

the compounds with the lowest $\kappa_{\ell}^{\text{TDEP}}$. Appendix: Table A.1 provides the numerical values of $\kappa_{\ell}^{\text{TDEP}}$ and $\kappa_{\ell}^{\text{AS}}$ at 500 K.

Table A.1

Calculated $\kappa_{\ell}^{\text{TDEP}}$ at 500 K for the HHs. The 32 compounds in the final test set are highlighted with bold text, and predictions made with the active sampling model (tier-1) are in parenthesis.

	κ_{ℓ} [W/mK]		κ_{ℓ} [W/mK]		κ_{ℓ} [W/mK]		κ_{ℓ} [W/mK]
LaPtSb	0.85	ZrPtGe	8.27	ZrCoBi	11.54 (10.41)	TaCoGe	15.47
LaRhTe	1.11 (1.79)	TiPdSn	8.42	NbIrSn	11.57 (11.13)	VRuAs	15.52
BaBiK	1.99	NbRuBi	8.70 (8.18)	HfCoAs	11.64	TiPdGe	15.78
CaPNa	2.58	TiPtSn	8.88	TiRhSb	11.72	GaNiNb	15.88 (14.65)
LiZnSb	3.76 (4.92)	HfIrAs	9.04	HfRhAs	11.76	NbOsAs	15.95
TaIrPb	5.20	ZrNiPb	9.10 (10.37)	TaIrSn	11.84	SiCoTa	16.28
TiPtPb	5.43	VOsSb	9.12	NbOsSb	11.86	VRhGe	16.57 (15.54)
BiPdSc	5.44 (5.92)	AlSiLi	9.18	ZrRhSb	11.89	VCoGe	16.58
BiNiY	5.46	ZrPtSn	9.20 (9.32)	HfCoBi	12.11 (10.20)	TaRuAs	16.60
HfPtPb	5.48	HfPdSn	9.32	NbCoSn	12.23	GaPtTa	16.75
HfPdPb	5.86	VRuSb	9.32	TaFeBi	12.40	VOsAs	16.82
TaOsBi	5.90	ZrPdSn	9.37	HfCoSb	12.48	TaRhGe	16.85
TiPdPb	5.94	AsNiSc	9.46	ZrNiGe	12.48	VRGe	16.95 (15.76)
TaRhPb	5.95	TaCoPb	9.66 (12.55)	NbRuSb	12.52	GeFeW	17.07
NbIrPb	6.07	NbRhSn	9.92 (11.44)	TiIrSb	12.73 (10.35)	TaFeSb	17.10 (14.85)
ZrPtPb	6.50 (6.48)	NbCoPb	9.95	TaRuSb	13.12 (11.83)	AlAuHf	17.11
HfIrBi	6.58	HfNiPb	9.96 (9.96)	TaCoSn	13.38	NbIrGe	17.21
ZrPdPb	6.59 (6.42)	HfPtSn	9.98	VFeSb	13.40	TiRhAs	17.52 (14.06)
TiIrBi	6.68	HfPdGe	10.28	HfIrSb	13.65	TaOsAs	18.09
ZrIrBi	6.88 (6.62)	TiNiSn	10.50 (13.83)	TeFeTi	13.82	TiCoBi	19.10
BiNiSc	6.97	ZrNiSn	10.84 (12.00)	TiCoSb	13.90 (13.87)	NbCoGe	19.37
HfRhBi	7.00	ZrIrSb	10.97	TiNiPb	13.91	TaIrGe	19.57
VIrSn	7.01	ZrPdGe	10.98	ZrCoAs	13.91 (13.97)	TiCoAs	19.92
TaRuBi	7.02 (7.58)	HfPtGe	11.08	TiIrAs	14.12	NbRuAs	20.21
VRhSn	7.11	TeRuZr	11.11	ZrCoSb	14.29	NbCoSi	20.31
ZrIrAs	7.20 (10.76)	HfNiSn	11.29	TiPtGe	14.34	NbRhGe	20.63
NbRhPb	7.21	VCoSn	11.30	ZrRhAs	14.54	NbFeAs	22.20
NbOsBi	7.26	TaRhSn	11.37	TiNiGe	14.78 (16.15)	VFeAs	22.60
ZrRhBi	7.33	NbFeBi	11.40	NbFeSb	15.00	LiBSi	23.45
AlGeLi	7.68	HfNiGe	11.41 (13.41)	TaFeAs	15.12 (16.37)		
TiRhBi	7.85	HfRhSb	11.49	TaOsSb	15.16 (12.16)		

Table 1 summarizes various performance metrics of the ML models. The baseline model predictions, $\log(\kappa_{\ell}^{\text{BL}})$, has a Spearman correlation of 0.79 with $\log(\kappa_{\ell}^{\text{TDEP}})$, which is larger than the correlation of Carrete et al. [18] of 0.74. Their model is based on fewer training samples, but more complex features. However, even though the Spearman correlation metric is fair and the low $\kappa_{\ell}^{\text{TDEP}}$ compounds do tend to be in the lower end of the spectrum of $\log(\kappa_{\ell}^{\text{BL}})$, the model fails to differentiate between the truly low $\kappa_{\ell}^{\text{TDEP}}$ compounds and the rest. The Spearman correlation of the active sampling model increases to 0.85. The superior ability of the active sampling model to predict properties of compounds with low $\kappa_{\ell}^{\text{TDEP}}$ results in improvement of the other performance metrics as well.

4.4. Machine learning at tier-0 level

The need for DFT-level input can be a drawback of using tier-1 features, as in some cases, experimental or calculated lattice constants are known while elastic tensors or bulk moduli are lacking. Thus, to uncover the potential of ML based on simpler features, we compare compound sampling and ML using tier-0 and tier-1 features. PCA-based active sampling with tier-0 features identifies the same three compounds as found earlier, as shown in Figs. 7 and 8(a). We allow for LiZnSb to be in the test set in this case, even though the drop-off in PC distance is less steep, as it allows for direct comparison of ML performance when using tier-0 and tier-1 features.

Fig. 8(b) compares the EFS selection frequencies of the active sampling models with tier-0 and tier-1 features. The average number of features chosen in the EFS for the active sampling model (tier-0) is 4.7, and increased selection frequencies are seen for χ_d and r_s .

Fig. 8(c) shows that the active sampling model (tier-0) identifies the lowest $\kappa_{\ell}^{\text{TDEP}}$ compound, but fails to differentiate the second and third lowest from the rest. This model has weaker predictivity overall compared to the active sampling model (tier-1), with an R2-score of 0.73 and Spearman correlation of 0.78. Despite that tier-0 features are used, the active sampling model (tier-0) outperforms the baseline model, underlining the importance of sample selection.

We also note that pre-sampling using PCA with tier-0 features could be used for pruning the test set. This can be done by excluding compounds that lie close to or within already sampled compound clusters in PC space. This would reduce the number of compounds in the test set, and thus reduce the computational resources needed for calculating the tier-1 DFT features.

While the performance gain when using more complex features and larger training sets has been demonstrated in earlier studies [20–23,60,73–75], this work demonstrates that using rather modest training set sizes and low feature complexity can give reliable predictions by adopting active sample selection.

On a cautionary note, the use of semi-random selection rather than truly random selection accentuates the performance gains when doing active sampling. We also find that only including one or two of the low lattice thermal conductivity compounds in the models significantly reduces performance compared to including all three. Performance with a truly random model would hence be sensitive to exactly which training samples are selected. In any case, a key advantage of PCA is that when used in the process to include additional HH compounds, we have a procedure to identify whether the properties of a given compound can be predicted reliably.

4.5. Comparison with experiments

Computed and predicted lattice thermal conductivities do not always agree perfectly with that of experiments. For comparison, the $\kappa_{\ell}^{\text{TDEP}}$ for NbCoSn and ZrNiSn of 12.2 W/mK and 10.8 W/mK are higher than lattice thermal conductivity measured in experiments: 7.0 W/mK [76] and 8.7 W/mK [77] for NbCoSn; 5.4 W/mK [78] and 6.1 W/mK [79] for ZrNiSn. Predictions made by the active sampling model (tier-1) for NbCoSn and ZrNiSn are 13.4 W/mK and 12.0 W/mK, respectively, when trained on calculated data. At the current level of theory, the difference between the ML predictions and TDEP lattice thermal conductivity is therefore much lower than that of experiment and theory. Sample dependent phonon scattering mechanisms due to

physical properties such as grain boundaries, intrinsic disorder, and antisite defects can drastically reduce lattice thermal conductivity [79–85], and can explain why experimentally measured lattice thermal conductivity is lower than predictions when computed lattice thermal conductivity only includes three-phonon and isotope scattering. Including such scattering mechanisms in the calculated lattice thermal conductivity can give values closer to that of experiment [81,86], which future studies should contemplate including in ML training for making models more representative of the lattice thermal conductivity of real-world samples.

5. Summary and conclusion

This study has explored strategies for using machine learning for finding low lattice thermal conductivity compounds using a limited number of training samples. Moreover, rather simple features were used, which can be found directly in material databases or computed straightforwardly. The exploration was made possible by computing lattice thermal conductivity with the temperature-dependent effective potential method for 122 half-Heusler compounds. We first demonstrated how a model based on a semi-random pool of materials (i.e. assumed “bad luck” in the training set) was unable to separate the truly low lattice thermal conductivity compounds in the test set from the rest. To improve the model, we used active sample selection based on principal component analysis. This approach suggested three compounds to be included in the training process. The subsequent inclusions resulted in a substantial improvement of model performance, in particular the ability to identify the remaining low lattice thermal conductivity compounds in the test set. Active sample selection without density functional theory-based features also identified necessary compounds to include in the model, but excluding the features in the model training resulted in weaker predictivity.

Our study demonstrates how active sampling can improve machine learning predictivity by accurately predicting properties of compounds dissimilar from the typical ones in a material class. More narrowly, we expect the procedure outlined here to be adopted to study broader classes of materials to systematically identify new low lattice thermal conductivity compounds.

CRediT authorship contribution statement

Rasmus Tranås: Conceptualization, Methodology, Software, Writing – original draft, Formal analysis, Investigation, Visualization. **Ole Martin Løvvik:** Conceptualization, Writing – review & feedback, Supervision, Funding acquisition. **Oliver Tomic:** Methodology, Writing – review & feedback, Supervision. **Kristian Berland:** Conceptualization, Methodology, Writing – review & feedback, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data forming the basis of this study is available from the authors upon reasonable request

Acknowledgments

The computations in this work were done on the high performance cluster Saga managed by UNINETT Sigma2. This work is part of the Allotherm project (Project No. 314778) supported by the Research Council of Norway. We further thank Øven Andreas Grimenes for helpful discussion and proofreading.

Appendix

Table A.1 shows $\kappa_{\ell}^{\text{TDEP}}$ for the HH compounds. Predictions made with the active sampling model (tier-1) are in the parenthesis.

References

- [1] G.J. Snyder, E.S. Toberer, Complex thermoelectric materials, *Nature Mater.* 7 (2) (2008) 105–114, <http://dx.doi.org/10.1038/nmat2090>.
- [2] D. Champier, Thermoelectric generators: A review of applications, *Energy Convers. Manag.* 140 (2017) 167–181, <http://dx.doi.org/10.1016/j.enconman.2017.02.070>.
- [3] J.J.G. Moreno, A review of recent progress in thermoelectric materials through computational methods, *Mater. Renew. Sustain. Energy* (2020) 22, <http://dx.doi.org/10.1007/s40243-020-00175-5>.
- [4] J. Wei, L. Yang, Z. Ma, P. Song, M. Zhang, J. Ma, F. Yang, X. Wang, Review of current high-ZT thermoelectric materials, *J. Mater. Sci.* 55 (27) (2020) 12642–12704, <http://dx.doi.org/10.1007/s10853-020-04949-0>.
- [5] J. Recatala-Gomez, A. Suwardi, I. Nandhakumar, A. Abutaha, K. Hippalgaonkar, Toward accelerated thermoelectric materials and process discovery, *ACS Appl. Energy Mater.* 3 (3) (2020) 2240–2257, <http://dx.doi.org/10.1021/acsaem.9b02222>.
- [6] P.P. Paskov, M. Slomski, J.H. Leach, J.F. Muth, T. Paskova, Effect of Si doping on the thermal conductivity of bulk GaN at elevated temperatures – theory and experiment, *AIP Adv.* 7 (9) (2017) 095302, <http://dx.doi.org/10.1063/1.4989626>.
- [7] S.Y. Kim, H.-S. Kim, K.H. Lee, H.-j. Cho, S.-s. Choo, S.-w. Hong, Y. Oh, Y. Yang, K. Lee, J.-H. Lim, S.-M. Choi, H.J. Park, W.H. Shin, S.-i. Kim, Influence of Pd doping on electrical and thermal properties of n-type $\text{Cu}_{0.008}\text{Bi}_2\text{Te}_3\text{Se}_{0.3}$ alloys, *Materials* 12 (24) (2019) 4080, <http://dx.doi.org/10.3390/ma12244080>.
- [8] K. Berland, N. Shulumba, O. Hellman, C. Persson, O.M. Løvvik, Thermoelectric transport trends in group 4 half-Heusler alloys, *J. Appl. Phys.* 126 (14) (2019) 145102, [arXiv:https://doi.org/10.1063/1.5117288](https://doi.org/10.1063/1.5117288).
- [9] R. Li, X. Li, L. Xi, J. Yang, D.J. Singh, W. Zhang, High-throughput screening for advanced thermoelectric materials: diamond-like ABX_3 compounds, *ACS Appl. Mater. Interfaces* 11 (28) (2019) 24859–24866, <http://dx.doi.org/10.1021/acsaami.9b01196>.
- [10] K. Choudhary, K.F. Garrity, F. Tavazza, Data-driven discovery of 3D and 2D thermoelectric materials, *J. Condens. Matter Phys.* 32 (47) (2020) 475501, <http://dx.doi.org/10.1088/1361-648X/aba06b>.
- [11] J. Liu, S. Han, G. Cao, Z. Zhou, C. Sheng, H. Liu, A high-throughput descriptor for prediction of lattice thermal conductivity of half-Heusler compounds, *J. Phys. D: Appl. Phys.* 53 (31) (2020) 315301, <http://dx.doi.org/10.1088/1361-6463/ab898e>.
- [12] J. He, M. Amsler, Y. Xia, S.S. Naghavi, V.I. Hegde, S. Hao, S. Goedecker, V. Ozoliņš, C. Wolverton, Ultralow thermal conductivity in full Heusler semiconductors, *Phys. Rev. Lett.* 117 (4) (2016) 046602, <http://dx.doi.org/10.1103/PhysRevLett.117.046602>.
- [13] P.R. Raghuvanshi, S. Mondal, A. Bhattacharya, A high throughput search for efficient thermoelectric half-Heusler compounds, *J. Mater. Chem. A* 8 (47) (2020) 25187–25197, <http://dx.doi.org/10.1039/D0A06810A>.
- [14] T.-T. Jia, Z. Feng, S. Guo, X. Zhang, Y. Zhang, Screening promising thermoelectric materials in binary chalcogenides through high-throughput computations, *ACS Appl. Mater. Interfaces* 12 (2020) 11852–11864, <http://dx.doi.org/10.1021/acsaami.9b23297>.
- [15] A. Togo, L. Chaput, I. Tanaka, Distributions of phonon lifetimes in Brillouin zones, *Phys. Rev. B* 91 (9) (2015) 094306, <http://dx.doi.org/10.1103/PhysRevB.91.094306>.
- [16] O. Hellman, I.A. Abrikosov, Temperature-dependent effective third-order interatomic force constants from first principles, *Phys. Rev. B* 88 (14) (2013) 144301, <http://dx.doi.org/10.1103/PhysRevB.88.144301>.
- [17] F. Zhou, W. Nielson, Y. Xia, V. Ozoliņš, Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations, *Phys. Rev. Lett.* 113 (18) (2014) 185501, <http://dx.doi.org/10.1103/PhysRevLett.113.185501>.
- [18] 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 (1) (2014) 011019, <http://dx.doi.org/10.1103/PhysRevX.4.011019>.
- [19] S. Ju, R. Yoshida, C. Liu, K. Hongo, T. Tadano, Exploring diamond-like lattice thermal conductivity crystals via feature-based transfer learning, 2019, p. 27, <http://dx.doi.org/10.1103/PhysRevMaterials.5.053801>.
- [20] 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, <http://dx.doi.org/10.1016/j.commatsci.2019.109155>.
- [21] X. Wang, S. Zeng, Z. Wang, J. Ni, Identification of crystalline materials with ultra-low thermal conductivity based on machine learning study, *J. Phys. Chem. C* 124 (16) (2020) 8488–8495, <http://dx.doi.org/10.1021/acs.jpcc.9b11610>.

- [22] C. Loftis, K. Yuan, Y. Zhao, M. Hu, J. Hu, Lattice thermal conductivity prediction using symbolic regression and machine learning, *J. Phys. Chem. A* 125 (1) (2021) 435–450, <http://dx.doi.org/10.1021/acs.jpca.0c08103>.
- [23] 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) 3559–3566, <http://dx.doi.org/10.1039/D1EE00442E>.
- [24] D. Visaria, A. Jain, Machine-learning-assisted space-transformation accelerates discovery of high thermal conductivity alloys, *Appl. Phys. Lett.* 117 (20) (2020) 202107, <http://dx.doi.org/10.1063/5.0028241>.
- [25] C.M. Collins, L.M. Daniels, Q. Gibson, M.W. Gaultois, M. Moran, R. Feetham, M.J. Pitcher, M.S. Dyer, C. Delacotte, M. Zanella, C.A. Murray, G. Glodan, O. Pérez, D. Pelloquin, T.D. Manning, J. Alaria, G.R. Darling, J.B. Claridge, M.J. Rosseinsky, Discovery of a low thermal conductivity oxide guided by probe structure prediction and machine learning, *Angew. Chem. Int. Ed.* (2021) <http://dx.doi.org/10.1002/anie.202102073>.
- [26] M.W. Gaultois, A.O. Olynyk, A. Mar, T.D. Sparks, G.J. Mulholland, B. Meredig, A recommendation engine for suggesting unexpected thermoelectric chemistries, *APL Mater.* 4 (5) (2016) 052313, <http://dx.doi.org/10.1063/1.4952607>, arXiv: 1502.07635.
- [27] F. Casper, T. Graf, S. Chadov, B. Balke, C. Felser, Half-Heusler compounds: novel materials for energy and spintronic applications, *Semicond. Sci. Technol.* 27 (6) (2012) 063001, <http://dx.doi.org/10.1088/0268-1242/27/6/063001>.
- [28] J.-W.G. Bos, R.A. Downie, Half-Heusler thermoelectrics: A complex class of materials, *J. Condens. Matter Phys.* 26 (43) (2014) 433201, <http://dx.doi.org/10.1088/0953-8984/26/43/433201>.
- [29] T. Zhu, C. Fu, H. Xie, Y. Liu, X. Zhao, High efficiency half-Heusler thermoelectric materials for energy harvesting, *Adv. Energy Mater.* 5 (19) (2015) 1500588, <http://dx.doi.org/10.1002/aenm.201500588>.
- [30] B. Yuan, B. Wang, L. Huang, X. Lei, L. Zhao, C. Wang, Q. Zhang, Effects of Sb substitution by Sn on the thermoelectric properties of ZrCoSb, *J. Electron. Mater.* 46 (5) (2017) 3076–3082, <http://dx.doi.org/10.1007/s11664-016-5168-z>.
- [31] A. Hori, S. Minami, M. Saito, F. Ishii, First-principles calculation of lattice thermal conductivity and thermoelectric figure of merit in ferromagnetic half-Heusler alloy CoMnSb, *Appl. Phys. Lett.* 116 (24) (2020) 242408, <http://dx.doi.org/10.1063/1.5143038>.
- [32] N.S. Chauhan, P.R. Raghuvanshi, K. Tyagi, K.K. Johari, L. Tyagi, B. Gahtori, S. Bathula, A. Bhattacharya, S.D. Mahanti, V.N. Singh, Y.V. Kolen'ko, A. Dhar, Defect engineering for enhancement of thermoelectric performance of (Zr, Hf)NiSn-based n-type half-Heusler alloys, *J. Phys. Chem. C* 124 (16) (2020) 8584–8593, <http://dx.doi.org/10.1021/acs.jpcc.0c00681>.
- [33] H. Zhu, J. Mao, Y. Li, J. Sun, Y. Wang, Q. Zhu, G. Li, Q. Song, J. Zhou, Y. Fu, R. He, T. Tong, Z. Liu, W. Ren, L. You, Z. Wang, J. Luo, A. Sotnikov, J. Bao, K. Nielsch, G. Chen, D.J. Singh, Z. Ren, Discovery of TaFeSb-based half-Heuslers with high thermoelectric performance, *Nature Commun.* 10 (1) (2019) 270, <http://dx.doi.org/10.1038/s41467-018-08223-5>.
- [34] H.-L. Sun, C.-L. Yang, M.-S. Wang, X.-G. Ma, Remarkably high thermoelectric efficiencies of the half-Heusler compounds BXGa (X=Be, Mg, and Ca), *ACS Appl. Mater. Interfaces* 12 (5) (2020) 5838–5846, <http://dx.doi.org/10.1021/acami.9b19198>.
- [35] D. Zhao, M. Zuo, Z. Wang, X. Teng, H. Geng, Synthesis and thermoelectric properties of tantalum-doped ZrNiSn half-Heusler alloys, *Funct. Mater. Lett.* 07 (03) (2014) 1450032, <http://dx.doi.org/10.1142/S1793604714500325>.
- [36] J. Zhou, H. Zhu, T.-H. Liu, Q. Song, R. He, J. Mao, Z. Liu, W. Ren, B. Liao, D.J. Singh, Z. Ren, G. Chen, Large thermoelectric power factor from crystal symmetry-protected non-bonding orbital in half-Heuslers, *Nature Commun.* 9 (2018) <http://dx.doi.org/10.1038/s41467-018-03866-w>.
- [37] Z. Feng, Y. Fu, Y. Zhang, D.J. Singh, Characterization of rattling in relation to thermal conductivity: ordered half-Heusler semiconductors, *Phys. Rev. B* 101 (6) (2020) 064301, <http://dx.doi.org/10.1103/PhysRevB.101.064301>, arXiv: 2001.08029.
- [38] Q.Y. Xue, H.J. Liu, D.D. Fan, L. Cheng, B.Y. Zhao, J. Shi, LaPtSb: A half-Heusler compound with high thermoelectric performance, *Phys. Chem. Chem. Phys.* 18 (27) (2016) 17912–17916, <http://dx.doi.org/10.1039/C6CP03211G>.
- [39] S.H. Han, Z.Z. Zhou, C.Y. Sheng, J.H. Liu, L. Wang, H.M. Yuan, H.J. Liu, High thermoelectric performance of half-Heusler compound BiBaK with intrinsically low lattice thermal conductivity, *J. Condens. Matter Phys.* 32 (42) (2020) 425704, <http://dx.doi.org/10.1088/1361-648X/aba2e7>.
- [40] G. Samsonidze, B. Kozinsky, Accelerated screening of thermoelectric materials by first-principles computations of electron-phonon scattering, *Adv. Energy Mater.* 8 (20) (2018) 1800246, <http://dx.doi.org/10.1002/aenm.201800246>.
- [41] S. Zhang, B. Xu, Y. Lin, C. Nan, W. Liu, First-principles study of the layered thermoelectric material TiNbR, *RSC Adv.* 9 (23) (2019) 12886–12894, <http://dx.doi.org/10.1039/C9RA00247B>.
- [42] D. Sarkar, T. Ghosh, S. Roychowdhury, R. Arora, S. Sajan, G. Sheet, U.V. Waghmare, K. Biswas, Ferroelectric instability induced ultralow thermal conductivity and high thermoelectric performance in rhombohedral p-type GeSe crystal, *J. Am. Chem. Soc.* 142 (28) (2020) 12237–12244, <http://dx.doi.org/10.1021/jacs.0c03696>.
- [43] L.-D. Zhao, S.-H. Lo, Y. Zhang, H. Sun, G. Tan, C. Uher, C. Wolverton, V.P. Dravid, M.G. Kanatzidis, Ultralow thermal conductivity and high thermoelectric figure of merit in SnSe crystals, *Nature* 508 (7496) (2014) 373–377, <http://dx.doi.org/10.1038/nature13184>.
- [44] G. Kresse, J. Hafner, *Ab initio* molecular dynamics for liquid metals, *Phys. Rev. B* 47 (1) (1993) 558–561, <http://dx.doi.org/10.1103/PhysRevB.47.558>.
- [45] G. Kresse, J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* 6 (1) (1996) 15–50, [http://dx.doi.org/10.1016/0927-0256\(96\)00008-0](http://dx.doi.org/10.1016/0927-0256(96)00008-0).
- [46] G. Kresse, J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* 54 (16) (1996) 11169–11186, <http://dx.doi.org/10.1103/PhysRevB.54.11169>.
- [47] J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, L.A. Constantin, X. Zhou, K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, *Phys. Rev. Lett.* 100 (13) (2008) 136406, <http://dx.doi.org/10.1103/PhysRevLett.100.136406>.
- [48] G.I. Csonka, J.P. Perdew, A. Ruzsinszky, P.H. Phillipsen, S. Lebègue, J. Paier, O.A. Vydrov, J.G. Ángyán, Assessing the performance of recent density functionals for bulk solids, *Phys. Rev. B* 79 (15) (2009) 155107, <http://dx.doi.org/10.1103/PhysRevB.79.155107>.
- [49] O. Hellman, I.A. Abrikosov, S.I. Simak, Lattice dynamics of anharmonic solids from first principles, *Phys. Rev. B* 84 (18) (2011) 180301, <http://dx.doi.org/10.1103/PhysRevB.84.180301>.
- [50] S.-i. Tamura, Isotope scattering of dispersive phonons in Ge, *Phys. Rev. B* 27 (2) (1983) 858–866, <http://dx.doi.org/10.1103/PhysRevB.27.858>.
- [51] S.-i. Tamura, Isotope scattering of large-wave-vector phonons in GaAs and InSb: Deformation-dipole and overlap-shell models, *Phys. Rev. B* 30 (2) (1984) 849–854, <http://dx.doi.org/10.1103/PhysRevB.30.849>.
- [52] N. Shulumba, O. Hellman, A.J. Minnich, Intrinsic localized mode and low thermal conductivity of PbSe, *Phys. Rev. B* 95 (2017) 014302, <http://dx.doi.org/10.1103/PhysRevB.95.014302>.
- [53] O. Anderson, A simplified method for calculating the Debye temperature from elastic constants, *J. Phys. Chem. Solids* 24 (7) (1963) 909–917, [http://dx.doi.org/10.1016/0022-3697\(63\)90067-2](http://dx.doi.org/10.1016/0022-3697(63)90067-2).
- [54] L. Breiman, Random forests, *Mach. Learn.* 45 (1) (2001) 5–32, <http://dx.doi.org/10.1023/A:1010933404324>.
- [55] S. Raschka, V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow*, 2nd Edition, second ed., Packt Publishing, 2017.
- [56] H.K. Jabbar, R.Z. Khan, Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study), in: *Computer Science, Communication and Instrumentation Devices*, Research Publishing Services, 2014, pp. 163–172, <http://dx.doi.org/10.3850/978-981-09-5247-1-017>.
- [57] S. Raschka, MLxtend: Providing machine learning and data science utilities and extensions to python's scientific computing stack, *J. Open Source Softw.* 3 (24) (2018) 638, <http://dx.doi.org/10.21105/joss.00638>.
- [58] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, Scikit-learn: machine learning in python, *J. Mach. Learn. Res.* 12 (2011) 2825.
- [59] O. Tomic, T. Graff, K. Liland, T. Næs, HOGGORM: A python library for explorative multivariate statistics, *J. Open Source Softw.* 4 (39) (2019) 980, <http://dx.doi.org/10.21105/joss.00980>.
- [60] H. Miyazaki, T. Tamura, M. Mikami, K. Watanabe, N. Ide, O.M. Ozkendir, N. Toichi, Machine learning based prediction of lattice thermal conductivity for half-Heusler compounds using atomic information, *Sci. Rep.* (2021) <http://dx.doi.org/10.1038/s41598-021-92030-4>.
- [61] H. Liang, V. Stanev, A.G. Kusne, I. Takeuchi, CRYSPNet: Crystal structure predictions via neural networks, *Phys. Rev. Mater.* 4 (2020) 123802, <http://dx.doi.org/10.1103/PhysRevMaterials.4.123802>.
- [62] Y. Li, W. Yang, R. Dong, J. Hu, MLatticeABC: Generic lattice constant prediction of crystal materials using machine learning, *ACS Omega* 6 (2021) 11585–11594, <http://dx.doi.org/10.1021/acsomega.1c00781>.
- [63] J. Mejia, T. Coplen, M. Berglund, W. Brand, P. De Bièvre, M. Gröning, N. Holden, J. Irrgeher, R. Loss, T. Walczyk, T. Prohaska, Atomic weights of the elements 2013 (IUPAC technical report), *Pure Appl. Chem.* 88 (2016) 265–291, <http://dx.doi.org/10.1515/pac-2015-0305>.
- [64] L.C. Allen, Electronegativity is the average one-electron energy of the valence-shell electrons in ground-state free atoms, *J. Am. Chem. Soc.* 111 (1989) 9003–9014, <http://dx.doi.org/10.1021/JA00207A003>.
- [65] B. Cordero, V. Gómez, A. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán, S. Alvarez, Covalent radii revisited, *Dalton Trans.* 21 (2008) 2832–2838, <http://dx.doi.org/10.1039/b801115j>.
- [66] G.A. Slack, Nonmetallic crystals with high thermal conductivity, *J. Phys. Chem. Solids* 34 (2) (1973) 15, [http://dx.doi.org/10.1016/0022-3697\(73\)90092-9](http://dx.doi.org/10.1016/0022-3697(73)90092-9).
- [67] T. Jia, G. Chen, Y. Zhang, Lattice thermal conductivity evaluated using elastic properties, *Phys. Rev. B* 95 (15) (2017) 155206, <http://dx.doi.org/10.1103/PhysRevB.95.155206>.
- [68] S. Ju, J. Shiomi, Materials informatics for heat transfer: recent progresses and perspectives, *Nanoscale Microscale Thermophys. Eng.* 23 (2) (2019) 157–172, <http://dx.doi.org/10.1080/15567265.2019.1576816>.

- [69] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, Commentary: The materials project: A materials genome approach to accelerating materials innovation, *APL Mater.* 1 (1) (2013) 011002, <http://dx.doi.org/10.1063/1.4812323>.
- [70] R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T.O.L. Sunde, D. Chon, K.R. Poeppelmeier, A. Zunger, Prediction and accelerated laboratory discovery of previously unknown 18-electron ABX compounds, *Nature Chem.* 7 (4) (2015) 308–316, <http://dx.doi.org/10.1038/nchem.2207>.
- [71] M.W. Gaultois, T.D. Sparks, How much improvement in thermoelectric performance can come from reducing thermal conductivity? *Appl. Phys. Lett.* 104 (11) (2014) 113906, <http://dx.doi.org/10.1063/1.4869232>.
- [72] J. An, A. Subedi, D. Singh, Ab initio phonon dispersions for PbTe, *Solid State Commun.* 148 (9–10) (2008) 417–419, <http://dx.doi.org/10.1016/j.ssc.2008.09.027>.
- [73] R. Juneja, G. Yumnam, S. Satsangi, A.K. Singh, Coupling the high-throughput property map to machine learning for predicting lattice thermal conductivity, *Chem. Mater.* 31 (14) (2019) 5145–5151, <http://dx.doi.org/10.1021/acs.chemmater.9b01046>.
- [74] R. Juneja, A.K. Singh, Guided patchwork kriging to develop highly transferable thermal conductivity prediction models, *J. Phys. Mater.* 3 (2) (2020) 024006, <http://dx.doi.org/10.1088/2515-7639/ab78f2>.
- [75] A. Tewari, S. Dixit, N. Sahni, S.P. Bordas, Machine learning approaches to identify and design low thermal conductivity oxides for thermoelectric applications, *Data-Centric Eng.* 1 (2020) e8, <http://dx.doi.org/10.1017/dce.2020.7>.
- [76] R. He, L. Huang, Y. Wang, G. Samsonidze, B. Kozinsky, Q. Zhang, Z. Ren, Enhanced thermoelectric properties of n-type NbCoSn half-Heusler by improving phase purity, *APL Mater.* 4 (2016) 104804, <http://dx.doi.org/10.1063/1.4952994>.
- [77] R. Yan, W. Xie, B. Balke, G. Chen, A. Weidenkaff, Realizing p-type NbCoSn half-Heusler compounds with enhanced thermoelectric performance via Sc substitution, *Sci. Technol. Adv. Mater.* 21 (2020) <http://dx.doi.org/10.1080/14686996.2020.1726715>.
- [78] A. Bahrami, P. Ying, U. Wolff, N.P. Rodriguez, G. Schierning, K. Nielsch, R. He, Reduced lattice thermal conductivity for half-Heusler ZrNiSn through cryogenic mechanical alloying, *ACS Appl. Mater. Interfaces* 13 (32) (2021) 38561–38568, <http://dx.doi.org/10.1021/acsami.1c05639>.
- [79] H. Xie, H. Wang, C. Fu, Y. Liu, G. Snyder, X. Zhao, T. Zhu, The intrinsic disorder related alloy scattering in ZrNiSn half-Heusler thermoelectric materials, *Sci. Rep.* 4 (2014) 6888, <http://dx.doi.org/10.1038/srep06888>.
- [80] M. Schrade, K. Berland, S. Eliassen, M. Guzik, C. Echevarria-Bonet, M. Sørby, P. Jenus, B. Hauback, R. Tofan, A. Gunnæs, C. Persson, O. Løvvik, T. Finstad, The role of grain boundary scattering in reducing the thermal conductivity of polycrystalline XNiSn (X=Hf, Zr, Ti) half-Heusler alloys, *Sci. Rep.* 7 (2017) <http://dx.doi.org/10.1038/s41598-017-14013-8>.
- [81] S.N. Eliassen, A. Katre, G.K.H. Madsen, C. Persson, O.M. Løvvik, K. Berland, Lattice thermal conductivity of $Ti_xZr_{1-x}NiSn$ half-Heusler alloys calculated from first principles: Key role of nature of phonon modes, *Phys. Rev. B* 95 (2017) 045202, <http://dx.doi.org/10.1103/PhysRevB.95.045202>.
- [82] A. Katre, J. Carrete, N. Mingo, Unraveling the dominant phonon scattering mechanism in thermoelectric compound ZrNiSn, *J. Mater. Chem. A* 4 (2016) <http://dx.doi.org/10.1039/C6TA05868J>.
- [83] J. Carrete Montaña, N. Mingo, S. Wang, S. Curtarolo, Nanograined half-Heusler semiconductors as advanced thermoelectrics: An ab initio high-throughput statistical study, *Adv. Funct. Mater.* 24 (2014) <http://dx.doi.org/10.1002/adfm.201401201>.
- [84] H. Hazama, M. Matsubara, R. Asahi, T. Takeuchi, Improvement of thermoelectric properties for half-Heusler TiNiSn by interstitial Ni defects, *J. Appl. Phys.* 110 (2011) <http://dx.doi.org/10.1063/1.3633518>.
- [85] H. Miyazaki, O.M. Ozkendir, S. Günaydin, K. Watanabe, K. Soda, Y. Nishino, Probing local distortion around structural defects in half-Heusler thermoelectric NiZrSn alloy, *Sci. Rep.* 10 (2020) 19820, <http://dx.doi.org/10.1038/s41598-020-76554-9>.
- [86] H. Geng, X. Meng, H. Zhang, J. Zhang, Lattice thermal conductivity of nanograined half-Heusler solid solutions, *Appl. Phys. Lett.* 104 (20) (2014) 202104, <http://dx.doi.org/10.1063/1.4879248>.

Paper II

Attaining Low Lattice Thermal Conductivity in Half-Heusler Sublattice Solid Solutions: Which Substitution Site Is Most Effective?

R. Tranås, O. M. Løvvik, and K. Berland

PUBLISHED. *Electron. Mater.*, 3(1), 1-14, (2022)





Article

Attaining Low Lattice Thermal Conductivity in Half-Heusler Sublattice Solid Solutions: Which Substitution Site Is Most Effective?

Rasmus Tranås ¹, Ole Martin Løvrik ^{2,3} and Kristian Berland ^{1,*}

¹ Department of Mechanical Engineering and Technology Management, Norwegian University of Life Sciences, NO-1432 Ås, Norway; rasmus.andre.tranas@nmbu.no

² Centre for Materials Science and Nanotechnology, Department of Physics, University of Oslo, NO-0316 Oslo, Norway; olemartin.lovvik@sintef.no

³ SINTEF Sustainable Energy Technology, NO-0314 Oslo, Norway

* Correspondence: kristian.berland@nmbu.no; Tel.: +47-456-792-996

† Current address: Drøbakveien 31, 1432 Ås, Norway.

Abstract: Low thermal conductivity is an important materials property for thermoelectricity. The lattice thermal conductivity (LTC) can be reduced by introducing sublattice disorder through partial isovalent substitution. Yet, large-scale screening of materials has seldom taken this opportunity into account. The present study aims to investigate the effect of partial sublattice substitution on the LTC. The study relies on the temperature-dependent effective potential method based on forces obtained from density functional theory. Solid solutions are simulated within a virtual crystal approximation, and the effect of grain-boundary scattering is also included. This is done to systematically probe the effect of sublattice substitution on the LTC of 122 half-Heusler compounds. It is found that substitution on the three different crystallographic sites leads to a reduction of the LTC that varies significantly both between the sites and between the different compounds. Nevertheless, some common criteria are identified as most efficient for reduction of the LTC: The mass contrast should be large within the parent compound, and substitution should be performed on the heaviest atoms. It is also found that the combined effect of sublattice substitution and grain-boundary scattering can lead to a drastic reduction of the LTC. The lowest LTC of the current set of half-Heusler compounds is around 2 W/Km at 300 K for two of the parent compounds. Four additional compounds can reach similarly low LTC with the combined effect of sublattice disorder and grain boundaries. Two of these four compounds have an intrinsic LTC above ~15 W/Km, underlining that materials with high intrinsic LTC could still be viable for thermoelectric applications.

Keywords: half-Heusler; lattice thermal conductivity; alloying; optimal substitution site; density functional theory; temperature-dependent effective potential



Citation: Tranås, R.; Løvrik, O.M.; Berland, K. Attaining Low Lattice Thermal Conductivity in Half-Heusler Sublattice Solid Solutions: Which Substitution Site Is Most Effective?. *Electron. Mater.* **2022**, *3*, 1–14. <https://doi.org/10.3390/electronicmat3010001>

Academic Editor: Ichiro Terasaki

Received: 15 November 2021

Accepted: 16 December 2021

Published: 5 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ability to convert excess heat into electricity and vice versa makes thermoelectricity interesting for a range of niche applications requiring local cooling or electricity generation. More effective thermoelectric materials could increase the applicability of thermoelectricity and thereby contribute to reducing energy consumption and carbon emissions [1,2]. The figure-of-merit of a thermoelectric material is at an operational temperature T given by

$$ZT = \frac{\sigma S^2 T}{\kappa_e + \kappa_\ell}, \quad (1)$$

where S is the Seebeck coefficient, σ is the electrical conductivity, and κ_e and κ_ℓ are the electronic and lattice thermal conductivity (LTC), respectively. The electronic contributions

to ZT exhibit conflicting dependency of the charge carrier concentration, requiring optimization of the doping level [3–6], though there has been significant progress in recent years [7–10].

Alloying, i.e., formation of solid solutions, is a widely used approach to modify materials, including their thermoelectric, optoelectronic [11,12], and electrolytic properties [13,14]. It can also be used to modify band gaps, defect formation densities, and structural phase transitions [15–19]. Introducing additional elements can, however, complicate materials synthesis e.g., by undesirable formation of secondary phases [20–26]. Still, much progress has been made in ensuring single-phase alloys for certain material classes [27] including half-Heuslers [28–30].

Half-Heusler (HH) compounds are ternary compounds with an XYZ composition, where the XZ sublattice forms a rocksalt structure and the YZ sublattice represents a zinc-blende structure [31]. The thermoelectric properties of HHs have been intensively studied [32–44]. While their electronic properties tend to be beneficial, they typically suffer from high LTC [45].

For optimizing thermoelectric properties, many studies have focused on identifying alloys with band structures that can provide a high power factor $\mathcal{P} = \sigma S^2$ without increasing κ_ℓ too much [46–48]. For such materials, the problem is reduced to that of lowering κ_ℓ while maintaining beneficial electronic properties [6]. An efficient strategy to reduce the LTC is to introduce sublattice disorder, for instance through partial isovalent atomic substitutions on one or more of the sublattices of the compound [49–54]. Such substitutions can be effective as they often have a limited effect on the electronic mobility [55–58] and hence the attainable \mathcal{P} .

Several studies have explored partial substitution on the X-site of HHs, e.g., replacing a fraction of Zr in ZrNiSn by Ti or Hf [59–62], or replacing a fraction of Nb in NbCoSn by Sc or Ti [63,64]. Z-site substitution has also been investigated in ZrCoSb, replacing Sb by Sn [65,66]. It would be useful to establish a systematic overview of the effectiveness of substitutions at different sites at reducing the LTC. An indicator of this can be provided by the magnitude of the mass-variance parameter [67], given as

$$g_i = \sum_j c_i^j \left(\frac{m_i^j - \bar{m}_i}{\bar{m}_i} \right)^2. \quad (2)$$

Here, c_i^j is the concentration of atom j substituting atom i , m_i^j is the mass of the substituent atom, and $\bar{m}_i = \sum_j c_i^j m_i^j$ is the average atomic mass. However, the nature of the phonon modes also plays a significant role, as exemplified for $\text{Ti}_x\text{Zr}_y\text{Hf}_{1-x-y}\text{NiSn}$ in Ref. [61]. In this paper, we use first-principles calculations to assess the change in LTC when introducing sublattice disorder on each of the three crystallographic sites of 122 HHs.

2. Methods

2.1. Thermal Transport and Phonon Scattering Mechanisms

The LTC is calculated with the temperature-dependent effective potential (TDEP) approach [68,69], in which the finite-temperature second- and third-order force constants are extracted based on atomic displacements and forces. Within phonon Boltzmann transport theory employing the relaxation-time approximation, the LTC is given by

$$\kappa_\ell = \frac{1}{V} \sum_\lambda c_\lambda v_\lambda^2 \tau_\lambda^{\text{tot}}, \quad (3)$$

where $\lambda = (\mathbf{q}s)$ indicates the wave vector \mathbf{q} and the phonon mode s . Further, c_λ is the phonon heat capacity, v_λ the phonon group velocity, and $\tau_\lambda^{\text{tot}}$ the total phonon relaxation time.

The scattering rate from three-phonon scattering is given by

$$\frac{1}{\tau_\lambda} = \frac{\hbar\pi}{8} \sum_{\lambda'\lambda''} |\Phi_{\lambda\lambda'\lambda''}|^2 [(n_{\lambda'} + n_{\lambda''} + 1)\delta(\omega_\lambda - \omega_{\lambda'} - \omega_{\lambda''}) + 2(n_{\lambda'} - n_{\lambda''})\delta(\omega_\lambda - \omega_{\lambda'} + \omega_{\lambda''})]. \quad (4)$$

Here, n is the equilibrium phonon distribution function, and $\Phi_{\lambda\lambda'\lambda''}$ are the three-phonon matrix elements,

$$\Phi_{\lambda\lambda'\lambda''} = \sum_{ijk} \sum_{\alpha\beta\gamma} \frac{\epsilon_{\alpha i}^\lambda \epsilon_{\beta j}^{\lambda'} \epsilon_{\gamma k}^{\lambda''}}{\sqrt{m_i m_j m_k} \sqrt{\omega_\lambda \omega_{\lambda'} \omega_{\lambda''}}} \times \Phi_{ijk}^{\alpha\beta\gamma} e^{i(\mathbf{q}\mathbf{r}_i + \mathbf{q}'\mathbf{r}_j + \mathbf{q}''\mathbf{r}_k)}. \quad (5)$$

Here, ijk are indices of atoms, $\alpha\beta\gamma$ are Cartesian coordinates, m_i is the mass of atom i , $\epsilon_{\alpha i}^\lambda$ is component α of the eigenvector for mode λ and atom i , \mathbf{r}_i is the lattice vector of atom i , and $\Phi_{ijk}^{\alpha\beta\gamma}$ are the third-order force constants. The mass-disorder (md) scattering rate is given by [67]

$$\frac{1}{\tau_\lambda^{\text{md}}} = \frac{\pi}{2} \sum_{\lambda'} \omega_\lambda \omega_{\lambda'} \sum_i g_i |\epsilon_{\lambda'}^{\dagger i} \cdot \epsilon_{\lambda}^i|^2 \delta(\omega_\lambda - \omega_{\lambda'}), \quad (6)$$

where $\epsilon_{\lambda'}^i$ is the polarization vector and g_i is the mass-variance parameter. For the unmixed compounds, τ_λ^{md} only contains the effect of natural isotopes, while for mixed compounds it also includes the distribution of isovalent mixing. These effects are both captured in g_i , but it should be noted that the natural isotope contribution is considerably smaller than the contribution of the mixing distribution [61,67].

We also investigate the effect of grain-boundary (gb) scattering. The scattering rate of grain boundaries is given by [70]

$$\frac{1}{\tau_\lambda^{\text{gb}}} = \frac{v_\lambda}{d}, \quad (7)$$

where v_λ is the phonon velocity and d is the characteristic domain (grain) size. We denote the intrinsic LTC arising from the combined effect of three-phonon scattering and isotope scattering as κ_ℓ^{mt} , with the additional effect of md-scattering as κ_ℓ^{md} , and with gb-scattering added on top of that as $\kappa_\ell^{\text{md,gb}}$.

2.2. Alloying Scheme

The 122 HHs in this study are based on the 74 stable compounds found by Feng et al. [71], as well as the additional 48 ones studied in Ref. [72]. Sublattice disorder is introduced by partially substituting with the element of the same group in the next period except for period 6 elements, which are substituted with period 5 elements. One exception, Te, is substituted with Se to omit the radioactive Po. For sublattice solid solutions, consisting solely of compounds predicted to be stable by Feng et al. [71] which are listed in the supplementary material (SM), we linearly interpolate the lattice parameter and the force constants of the parent compound with those of the fully substituted compound. Throughout the study, we use a 10% mixing concentration. For instance, for $\text{Zr}_{0.9}\text{Hf}_{0.1}\text{NiSn}$, the X-site effective mass is equal to 90% Zr plus 10% Hf isotopic mass. Likewise, the lattice parameter and the force constants consist of 90% of those calculated for ZrNiSn and 10% of those for HfNiSn . For the remaining compounds, we use force constants and lattice parameter from the parent compounds to limit computational costs.

2.3. Computational Details

Density functional theory (DFT) calculations are done using the VASP [73–75] software package using the generalized gradient approximation (GGA) functional PBEsol [76,77]. The plane-wave cutoff is set to 600 eV. The electronic self-consistency cutoff is set at 10^{-6} eV. Atoms are relaxed until all forces fall below 1 meV/Å. Second- and third-order force

constants are extracted from fifty supercell configurations with $3 \times 3 \times 3$ repetitions of the primitive unit cell. For each configuration, atoms are displaced from equilibrium according to a canonical ensemble set to $T = 300$ K. Realistic atomic displacements are calculated by assuming that the zero-point energy of the phonons corresponds to the Debye temperature [78]. The Debye temperature is obtained from the elastic tensor [79], which is calculated with DFT using the frozen phonon approach. In the supercell calculations, we use a $3 \times 3 \times 3$ k -point sampling of the Brillouin zone. The elastic tensor calculations are done with the primitive unit cell and an $11 \times 11 \times 11$ k -point sampling.

When extracting the force constants, the cutoff radii of the third-order and second-order interactions are $\sim 2\%$ and $\sim 12\%$ larger than half the width of the supercell, respectively. These choices ensure that the force-constant matrices have equal sizes, allowing straightforward interpolation. The phonon-mode Brillouin zone integration in Equation (3) is sampled on a $35 \times 35 \times 35$ q -point mesh.

3. Results and Analysis

3.1. Lattice thermal conductivity with Mass-Disorder and Grain-Boundary Scattering

Figure 1a–c shows the LTC with intrinsic scattering, $\kappa_{\ell}^{\text{int}}$, and with the additional md-scattering included, $\kappa_{\ell}^{\text{md}}$, for the 122 HH compounds. Figure 1d plots $\kappa_{\ell}^{\text{int}}$ versus $\kappa_{\ell}^{\text{md}}$, where the color of the triangles indicates substitution site, displaying the reduction in LTC for each of the substitutions. On average, X-site substitution reduces the LTC by 43%, somewhat less than the $\sim 49\%$ reduction at Y- and Z-sites. There is a substantial variation in $\kappa_{\ell}^{\text{md}}$ when substituting the different sites, as shown by the large vertical spread of the triangles and the distributed bins of the histogram. To exemplify, partial X-site substitution of VIrGe (Figure 1c) reduces the LTC by 14%, while the reduction with Z-site substitution in LiBSi (Figure 1a) is 80%. The site-dependence can also differ greatly from one compound to the next. In the case of ZrRhSb and ZrPdPb (Figure 1c), the reduction is quite similar for the three different substitution sites, while for VFeSb (Figure 1c), X- and Z-site partial substitution leads to a reduction of 21% and 70%, respectively. In total, we find 91 compounds with more than a 25% difference between the largest and smallest $\kappa_{\ell}^{\text{md}}$ obtained for the different substitution sites.

Based on this data, it is interesting to formulate some basic guidelines on which substitution site to select to most efficiently reduce the LTC. One simple rule, which we previously noted as a trend in Ref. [80], is to substitute on the site hosting the heaviest atom. This rule gives the correct result for 102 out of the 122 compounds. Among the 20 compounds not following this rule, there are only 4 cases in which the $\kappa_{\ell}^{\text{md}}$ obtained with heaviest-site substitution and optimal-site substitution differ by more than 15%. These 4 counterexamples can be explained by the fact that the mass-variance parameter g_i of the optimal-site substitutions for these compounds is more than 50% larger than that of the heaviest-site substitution.

Another possible rule for attaining the largest LTC reduction could be to use the substitution with the highest g_i . In general, however, this turns out to be less efficient than substituting on sites with heavy atoms; often a high g_i coincides with lighter atoms, which are associated with high-frequency phonon contributions to heat transport [61]. This is exemplified by NbCoPb, which is optimally substituted on the Z-site. Given the relatively similar mass of Pb and Sn, Z-site substitution results in merely $g_Z = 0.02$; the lighter Co partially substituted by Rh yields, on the other hand, $g_Y = 0.04$, while Nb substituted by Ta corresponds to $g_X = 0.07$.

Figure 1e shows the effect of introducing gb-scattering on the LTC, including optimal-site substitutions. The relative drop in LTC is larger for compounds with high $\kappa_{\ell}^{\text{int}}$. For the 20 compounds with lowest $\kappa_{\ell}^{\text{int}}$, the reduction in LTC using 100 nm grains and substitution on the optimal site is on average 69%, while for the 20 compounds with the highest $\kappa_{\ell}^{\text{int}}$, the average reduction is 85%. When optimal-site substitution has been included, reducing the grain size from 100 nm to 50 nm only gives a modest reduction in LTC; the average drop compared to the intrinsic LTC changes from 79% to 83%.

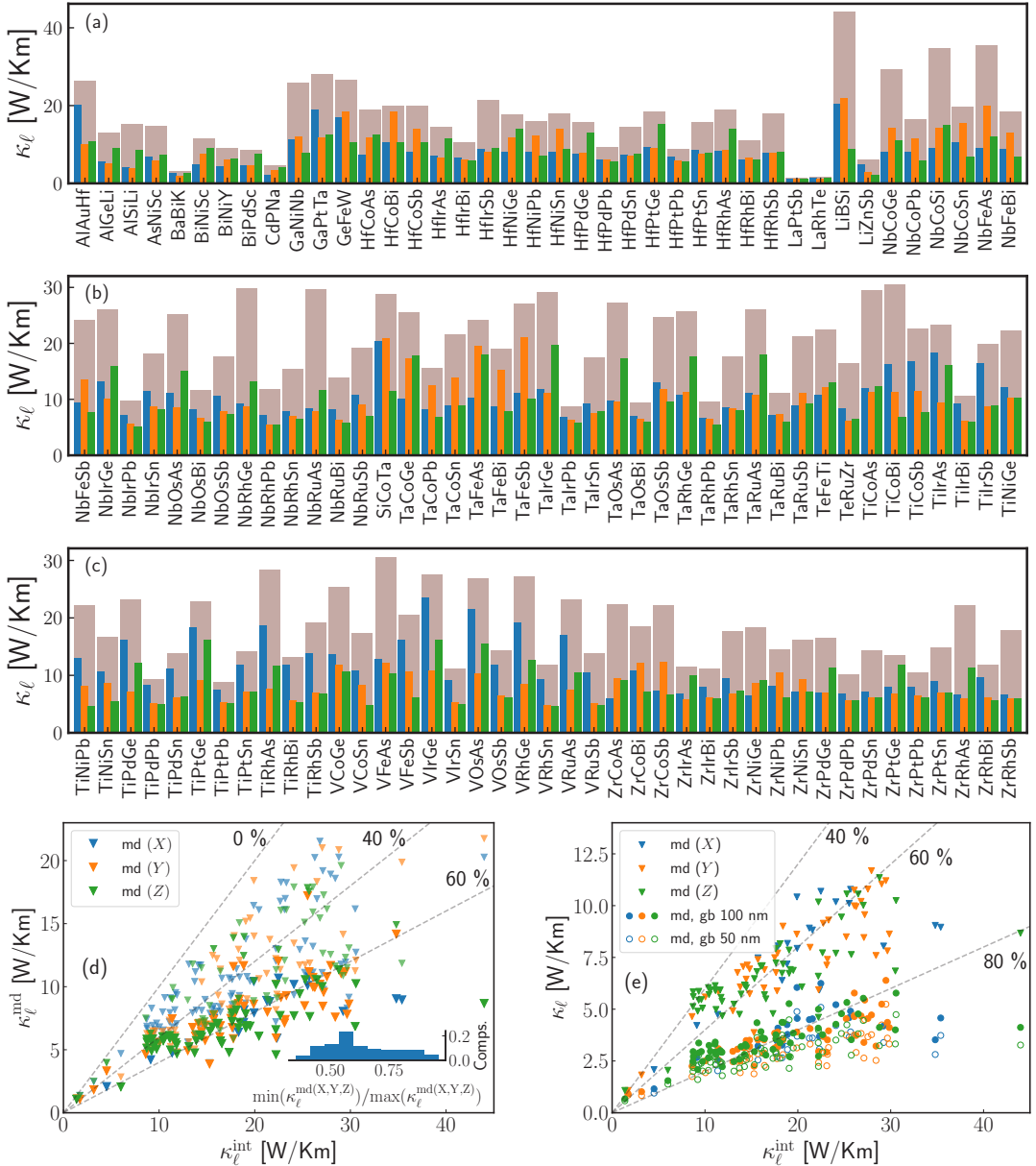


Figure 1. The (a–c) panels compare the LTC involving only intrinsic scattering κ_ℓ^{int} (grey bars) with that obtained from the addition of md-scattering on sites X, Y, and Z (blue, orange, green), i.e., $\kappa_\ell^{\text{md}(X)}$ at 300 K. Dots mark the optimal substitution site. Panel (d) plots κ_ℓ^{int} against κ_ℓ^{md} . The blue, orange, and green triangles indicate substitution on the X-, Y-, and Z-sites, respectively, with the one most effectively reducing the LTC highlighted. The inset histogram shows the distribution of the most effective substitution divided by the least effective one, $\min(\kappa_\ell^{\text{md}(X,Y,Z)})/\max(\kappa_\ell^{\text{md}(X,Y,Z)})$. In panel (e), only results for the most effective substitution site are shown; in addition filled (empty) discs indicate $\kappa_\ell^{\text{md,gb}}$ with gb size of 100 nm (50 nm).

Grain boundaries most efficiently scatter less energetic acoustic phonons [62,81], while mass disorder scatters more energetic phonons [61,82,83]. It is therefore interesting to assess the reduction achieved with gb-scattering without the inclusion of md-scattering. We find that introducing gb-scattering with grain sizes of 100 and 50 nm reduces LTC on average by 42% and 54%, respectively, see details in the SM. This reduction is hence more modest than that of 10% substitution on the optimal site.

3.2. Compounds with Low Lattice Thermal Conductivity

Certain compounds reach very low LTC with the introduction of md- and gb-scattering. Some of these are simply compounds that already have a low intrinsic LTC, such as LaPtSb, LaRhTe, BaBiK, and CdPNa. For these compounds, it is interesting to gauge whether additional scattering mechanisms provide a substantial reduction in the LTC. Others are compounds which exhibits a significant drop in LTC with the inclusion of additional scattering mechanisms. It is useful to identify the key characteristics of these compounds, for the sake of developing effective thermoelectric optimisation and material identification strategies.

In Figure 2, we compare the intrinsic LTC, $\kappa_{\ell}^{\text{int}}$, with the LTC where md-scattering has been included, $\kappa_{\ell}^{\text{md}}$, for the compounds with lowest $\kappa_{\ell}^{\text{md}}$ at temperatures 300 K, 500 K, and 700 K. As a reference, NbCoSn, ZrNiSn, and the average LTCs for all HHs are also included. For partial substitution on the optimal site (reducing LTC the most), we also display the effect of gb-scattering with a grain size of 100 nm.

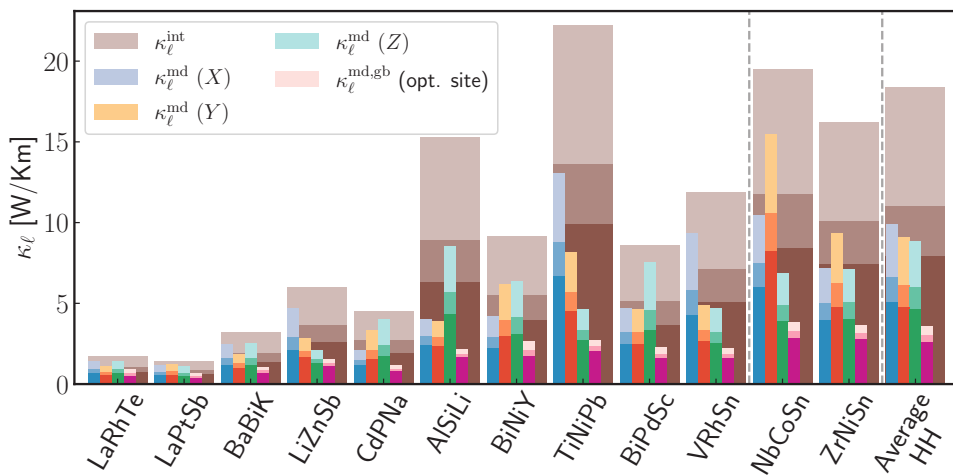


Figure 2. $\kappa_{\ell}^{\text{int}}$ and $\kappa_{\ell}^{\text{md}}$ for the 10 HH compounds with the lowest $\kappa_{\ell}^{\text{md}}$ at 300 K, compared with NbCoSn, ZrNiSn, and the average of all HH compounds as reference. Brown bars indicate the intrinsic LTC, while blue, orange, and green bars indicate the LTC with substitutions on X-, Y-, and Z-sites. Sub-bars of increasing color intensity indicate the LTC at 300 K, 500 K, and 700 K. Purple bars indicate $\kappa_{\ell}^{\text{md,gb}}$, where substitution is done on the site reducing LTC the most, including gb-scattering using a domain size of 100 nm.

For the compounds with the largest potential drop in LTC, there is a strong substitution site sensitivity. In general, we find that these compounds are characterized by a large mass difference between the heaviest and lightest atoms in the parent compound, with substitution on the heaviest atom being preferable. This can be exemplified by TiNiPb and AlSiLi, which exhibit the largest potential drop in LTC with the introduction of md-scattering; Pb is much heavier than Ti and Ni, and Al and Si are much heavier than Li. TiNiPb has a relatively high $\kappa_{\ell}^{\text{int}} = 22.2$ W/Km at 300 K, but Z-site substitution brings it

down to $\kappa_{\ell}^{\text{md}} = 4.6$ W/Km, whereas X-site substitution results in 13.0 W/Km and Y-site results in 8.2 W/Km. AlSiLi has the largest reduction with Y-site substitution, resulting in $\kappa_{\ell}^{\text{md}} = 3.9$ W/Km, quite similar to that of X-site substitution, which results in 4.0 W/Km, while Z-site substitution results in 8.5 W/Km.

For the compounds with lowest $\kappa_{\ell}^{\text{int}}$, i.e., LaPtSb, LaRhTe, BaBiK, and CdPNa, substitution on the optimal site reduces, on average, the LTC by 38%, while the average reduction for all the HHs is 58%. Thus, while the reduction achieved by introducing additional scattering mechanisms is lower for these compounds, it is still substantial. It is also interesting to compare LiZnSb and CdPNa. Both these compounds have a quite similar mass-variance parameter g_i at the optimal substitution site: $g_i = 0.05$ for substitution on the X-site in CdPNa and $g_i = 0.04$ for Z-site substitution in LiZnSb. Nonetheless, the effect of substitution is far stronger for LiZnSb. Similar to AlSiLi, LiZnSb is characterized by one of the atoms being much lighter than the other two.

When gb-scattering is introduced in addition to md-scattering, LiZnSb, AlSiLi, BiPdSc, and VRhSn achieve $\kappa_{\ell}^{\text{md,gb}} \leq 2.2$ W/Km, comparable to the values of the compounds with the lowest intrinsic LTC without such scattering mechanisms. In comparison, NbCoSn and ZrNiSn have $\kappa_{\ell}^{\text{md,gb}}$ of 3.9 W/Km and 3.6 W/Km, and the average for the HHs of this study is $\kappa_{\ell}^{\text{md,gb}} = 3.6$ W/Km when using optimal-site substitutions.

Figure 3 shows the change in $\kappa_{\ell}^{\text{md,gb}}$ with temperature. While the relative ordering of low to high $\kappa_{\ell}^{\text{md,gb}}$ compounds tends to be retained with temperature, there is some difference in the relative reduction with increasing temperatures for different compounds. Interestingly, the compounds having a low intrinsic LTC, LaPtSb, LaRhTe, BaBiK, and CdPNa, have a relatively large reduction in $\kappa_{\ell}^{\text{md,gb}}$ with temperature, on average 44% when going from 300 to 800 K. The average reduction for all HHs is only 32% over the same temperature span. The compounds that attain a low $\kappa_{\ell}^{\text{md,gb}}$, LiZnSb, AlSiLi, BiPdSc, VRhSn, BiNiY, and TiNiPb, have an average reduction of 33%, i.e., similar to that of the average for the HHs. Among these, the largest reduction is found for BiNiY, for which the LTC decreases by 40% when going from 300 to 800 K. In general $\kappa_{\ell}^{\text{md,gb}}$ shows a more modest reduction with temperature than $\kappa_{\ell}^{\text{int}}$, which on average is reduced by 62%. The slight differences between the temperature dependence from one compound to the next, on the other hand, almost vanish without the presence of additional scattering mechanisms.

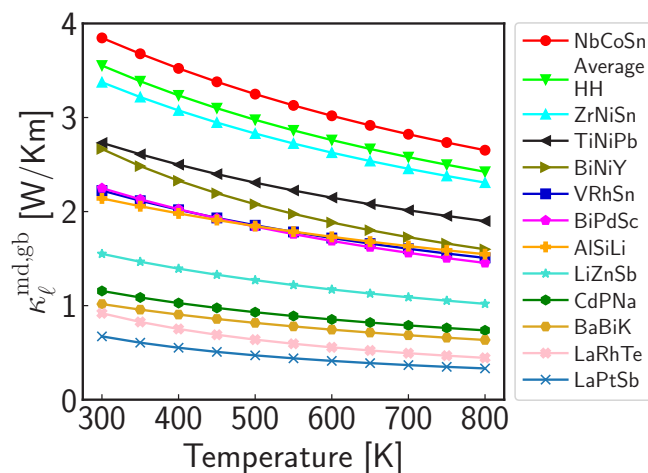


Figure 3. $\kappa_{\ell}^{\text{md,gb}}$ as a function of temperature, with substitution on the optimal site and a domain size of 100 nm. The temperature is indicated on the horizontal axis, and $\kappa_{\ell}^{\text{md,gb}}$ on the vertical axis.

4. Discussion

While our paper provides valuable guidance on the optimization of sublattice solid solutions for reducing the LTC, there are certain limitations of our analysis to bear in mind. First, the study is based on the virtual crystal approximation, in which the sublattice disorder is not accounted for explicitly. For a selection of the compounds, we simply use the average “virtual” mass on a site, while for others we also interpolate the force constants and lattice parameters, as discussed in Section 2.2. These limitations imply uncertainty in the calculated numbers.

Second, we neglect strain-induced scattering when the mass-disorder effect is considered. In a recent paper, Arrigoni et al. [84] showed that force-constant disorder can in some systems significantly reduce the predicted LTC. For $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ the reduction was $\sim 40\%$, while for $\text{Si}_{0.3}\text{Ge}_{0.7}$ it was less than 10%. The difference between the LTC obtained with interpolation of the mass, lattice parameter, and force constants compared with interpolation of only the mass provides some indication of the magnitude of the strain-induced scattering. This is because a large shift would imply that force constants of the two compounds are quite different, indicating large strain-induced scattering matrix elements. As detailed in the SM, we find that interpolating the lattice parameter and force constants changes the LTC on average by 6.6% for substitution on the X-site, 7.3% for the Y-site, and 7.2% for the Z-site. Thus, while this change is modest, it is not insubstantial; for some compounds, e.g., for Y-site substituted NbCoSn it is quite large, having a relative deviation of 23%. Nonetheless, md-scattering provides a rough lower bound of the scattering introduced through substitution. One could argue that a large mass-variance tends to correlate with variance in the effective size of the atoms and as such strain induced-scattering. Thus, one could hope that much of the strain-induced scattering would be described indirectly by the mass-variance. However, strain affects more atom sites than the mass-disorder, which is entirely local. The relative difference in the effect of substitution on different sites would therefore be somewhat affected by strain. The significance of this is difficult to evaluate, but we assume that the key trends identified here remain valid.

Third, we use a substitution concentration of 10% throughout this work. Increasing the concentration could further reduce LTC, but in terms of synthesis, higher concentrations can also lead to phase separation [66,85]. We do note, however, that phase separation could in some cases be beneficial for thermoelectric performance [86]. Sublattice miscibility is not considered in this paper, but several HH alloys demonstrate substitution concentrations higher than 10% [59,60,62,66,87,88]. However, some cases, such as 10% substitution of Li with Na in X-site substituted LiZnSb could be unfeasible.

Given the many approximations, we perform a rough comparison with experiment to obtain an indication of the accuracy of the predictions. In doing so, we first consider that the computed intrinsic LTC is often much larger than the experimental LTC, especially for unmixed compounds [61,62,72]. This overestimation can, at least partially, be explained by the many impurities and defects present in real-world samples [35,89–91]. To correct for this, we first compute an effective contribution to the thermal conductivity arising from md-scattering, based on Matthiessen’s rule: $1/\kappa'_\ell = 1/\kappa_\ell^{\text{md}} - 1/\kappa_\ell^{\text{int}}$. This expression can then be compared with a corresponding experimental quantity, $1/\kappa'_{\ell,\text{exp}} = 1/\kappa_{\ell,\text{exp}}^{\text{mixed}} - 1/\kappa_{\ell,\text{exp}}^{\text{parent}}$. An underlying assumption in this comparison is that the phonon scattering due to impurities and other defects not related to mixing is quite similar in the mixed ($\kappa_{\ell,\text{exp}}^{\text{mixed}}$) and pure compound ($\kappa_{\ell,\text{exp}}^{\text{parent}}$). Under this assumption, it is meaningful to compare $\kappa'_{\ell,\text{exp}}$ with the theoretical counterpart κ'_ℓ , allowing for a rough comparison of the computed effect of md-scattering with the mixing induced experimental effect. Comparing X-site substitution in ZrNiSn at 500 K, we obtain in this work $\kappa'_\ell = 9.9 \text{ W/Km}$, which is in reasonable agreement with the experimental value of $\kappa'_{\ell,\text{exp}} = 7.0 \text{ W/Km}$ for $\text{Zr}_{0.9}\text{Hf}_{0.1}\text{NiSn}$ [59]. Moreover, for X-site substitution in VFeSb at 300 K, we find $\kappa'_\ell = 21.8 \text{ W/Km}$, which can be compared with $\kappa'_{\ell,\text{exp}} = 18.7 \text{ W/Km}$ [92] for $\text{V}_{0.9}\text{Nb}_{0.1}\text{FeSb}$. These differences are sizeable but still

point to the general validity of using md-scattering as a rough indicator of the effect of mixing on the LTC.

We also present results both with and without gb-scattering. The gb-scattering model is highly idealized, assuming perfect scattering of long-wavelength acoustic phonons. Real-world materials have grains of different sizes and orientations which scatter phonons differently. Domain sizes of 50 and 100 nm are assumed here, while depending on the preparation method, experimentally average grain sizes have been reported in the 60 to 400 nm range [62,93,94]. Nonetheless, gb-scattering provides some indication of how grains and other macroscopic defects filter out the contributions of mostly long-wavelength phonons to the LTC.

On a final note, one must keep in mind that substitutions and grain boundaries also impact electron transport. The additional electron scattering introduced causes a reduction in the electron mobility and in turn the conductivity σ [60]. However, the subsequent reduction in ZT would be lessened by the simultaneous reduction in the electronic thermal conductivity κ_e . LaPtSb, LaRhTe, BaBiK, and CdPNa have $\kappa_e^{\text{md,gb}} \leq 1.2$ W/Km, comparable to well-known low LTC thermoelectric materials such as PbTe [95,96] and SnSe [10,97]. This value is similar to or even lower than κ_e in several doped HHs [80], e.g., $\kappa_e = 1.4$ W/Km for NbCo_{0.94}Ni_{0.06}Sn [98] and $\kappa_e = 1.5$ W/Km for ZrPtSn_{0.92}Sb_{0.08} [99]. Thus, with the identification of low LTC comparable to κ_e , emphasis can shift to design strategies that reduce κ_e .

5. Conclusions

This study has investigated the effect of isovalent sublattice-substitutions and grain boundaries on the lattice thermal conductivity (LTC) of 122 half-Heusler compounds. Calculations were done using density functional theory and the temperature-dependent effective potential method. We have shown that several compounds with relatively high LTC could be promising thermoelectric materials because of a substantial reduction in LTC with the introduction of additional phonon scattering mechanisms. In general, the largest reduction in LTC was achieved when substituting on the site hosting the heaviest atom rather than substituting on the site hosting the atom having the largest mass-variance parameter. We identified four compounds with quite large intrinsic LTC; LiZnSb, AlSiLi, BiPdSc, and VRhSn, for which the LTC fell below 2.2 W/Km at 300 K with the introduction of mass-disorder and grain-boundary scattering. For the four compounds with the lowest intrinsic LTC, we found that both mass-disorder and grain-boundary scattering can further reduce LTC, although the relative reduction is smaller, resulting in an LTC below 1.2 W/Km.

Our study highlights that the presence of elements with large masses can be beneficial for thermoelectric materials. Such elements both reduce phonon velocities and allow for large mass-disorder scattering. However, compounds with heavy atoms are typically associated with expensive and rare elements like La, Pt, and Pd, or toxic elements such as Pb and Te. This makes them less desirable for new large-scale thermoelectric applications. However, the low LTC of Ga-substituted AlSiLi and Bi-substituted LiZnSb illustrates that it can be sufficient that one or two of the atoms are substantially heavier than the others once mass-disorder and grain-boundary scattering effects are taken into account. It is interesting to note that the reason for the large drop in LTC seen for such compounds could be connected to the nature of phonon modes. The substantial mass difference in zinc-blende compounds has been linked to large gaps between acoustic and optical phonons and small three-phonon scattering phase-space volume [100]. Such compounds would therefore typically have intrinsic LTC on the higher end, and thus potentially a large reduction in LTC with the introduction of additional scattering mechanisms. The identification of compounds with a large drop in LTC with the introduction of additional scattering mechanisms highlights the importance of going beyond three-phonon scattering when screening materials for thermoelectric applications. The prospect of finding efficient thermoelectric materials that are also cheap, easy to manufacture, and non-toxic remains bright.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/electronicmat3010001/s1>.

Author Contributions: Conceptualization, K.B.; software, R.T.; formal analysis, R.T. and K.B.; investigation, R.T.; data curation, R.T.; writing—original draft preparation, R.T.; writing—review and editing, O.M.L. and K.B.; visualization, R.T. and K.B.; supervision, O.M.L. and K.B.; project administration, K.B.; funding acquisition, O.M.L. and K.B. All authors have read and agreed to the published version of the manuscript.

Funding: The computations in this work were funded by The Norwegian e-infrastructure for research and education, Sigma2, through grants No. NN9711K and NN2615K. This work is part of the Allotherm project (Project No. 314778) funded by the Research Council of Norway.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article or Supplementary Material.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HH	half-Heusler
DFT	density functional theory
LTC	lattice thermal conductivity
TDEP	temperature-dependent effective potential
md	mass-disorder
gb	grain-boundary
SM	supplementary material

References

- Ando Junior, O.; Maran, A.; Henao, N. A review of the development and applications of thermoelectric microgenerators for energy harvesting. *Renew. Sust. Energ. Rev.* **2018**, *91*, 376–393. [[CrossRef](#)]
- Champion, D. Thermoelectric generators: A review of applications. *Energy Convers. Manag.* **2017**, *140*, 167–181. [[CrossRef](#)]
- Snyder, G.; Toberer, E. Complex thermoelectric materials. *Nat. Mater.* **2008**, *7*, 105–114. [[CrossRef](#)] [[PubMed](#)]
- Zhu, T.; Liu, Y.; Fu, C.; Heremans, J.P.; Snyder, J.G.; Zhao, X. Compromise and Synergy in High-Efficiency Thermoelectric Materials. *Adv. Mater.* **2017**, *29*, 1605884. [[CrossRef](#)] [[PubMed](#)]
- Twaha, S.; Zhu, J.; Yan, Y.; Li, B. A comprehensive review of thermoelectric technology: Materials, applications, modelling and performance improvement. *Renew. Sust. Energ. Rev.* **2016**, *65*, 698–726. [[CrossRef](#)]
- Wongprakarn, S.; Pinitsoontorn, S.; Tanusilp, S.A.; Kurosaki, K. Enhancing thermoelectric properties of p-type SiGe alloy through optimization of carrier concentration and processing parameters. *Mater. Sci. Semicond. Process* **2018**, *88*, 239–249. [[CrossRef](#)]
- Yang, L.; Chen, Z.G.; Dargusch, M.; Zou, J. High performance thermoelectric materials: Progress and their applications. *Adv. Energy Mater.* **2017**, *8*, 1701797. [[CrossRef](#)]
- Gutierrez Moreno, J.; Cao, J.; Fronzi, M.; Assadi, M.H.N. A review of recent progress in thermoelectric materials through computational methods. *Mater. Renew.* **2020**, *9*, 16. [[CrossRef](#)]
- Gangjian, T.; Shi, F.; Hao, S.; Zhao, L.; Chi, H.; Zhang, X.; Uher, C.; Wolverton, C.; Dravid, V.; Kanatzidis, M. Non-equilibrium processing leads to record high thermoelectric figure of merit in PbTe–SrTe. *Nat. Commun.* **2016**, *7*, 12167. [[CrossRef](#)]
- Zhou, C.; Lee, Y.K.; Yu, Y.; Byun, S.; Luo, Z.Z.; Lee, H.; Ge, B.; Lee, Y.L.; Chen, X.; Lee, J.Y.; et al. Polycrystalline SnSe with a thermoelectric figure of merit greater than the single crystal. *Nat. Mater.* **2021**, *20*. [[CrossRef](#)]
- Wheatley, R.; Roble, M.; Gence, L.; Acuña, C.; Rojas-Aedo, R.; Hidalgo-Rojas, D.; Guzman-De La Cerda, D.; Vojkovic, S.; Seifert, B.; Wallentowitz, S.; et al. Structural, optoelectronic and photo-thermoelectric properties of crystalline alloy $\text{CuAl}_x\text{Fe}_{1-x}\text{O}_2$ delafossite oxide materials. *J. Alloys Compd.* **2021**, *857*, 157613. [[CrossRef](#)]
- Rajagopal, A.; Stoddard, R.J.; Hillhouse, H.W.; Jen, A.K.Y. On understanding bandgap bowing and optoelectronic quality in Pb–Sn alloy hybrid perovskites. *J. Mater. Chem. A* **2019**, *7*, 16285–16293. [[CrossRef](#)]
- Nikam, R.; Kwak, M.; Lee, J.; Rajput, K.; Banerjee, W.; Hwang, H. Near ideal synaptic functionalities in Li ion synaptic transistor using $\text{Li}_3\text{PO}_x\text{Se}_x$ electrolyte with high ionic conductivity. *Sci. Rep.* **2019**, *9*, 18883. [[CrossRef](#)] [[PubMed](#)]
- Zevgolis, A.; Wood, B.C.; Mehmedović, Z.; Hall, A.T.; Alves, T.C.; Adelstein, N. Alloying effects on superionic conductivity in lithium indium halides for all-solid-state batteries. *APL Mater.* **2018**, *6*, 047903. [[CrossRef](#)]

15. Haseman, M.S.; Karim, M.R.; Ramdin, D.; Noesges, B.A.; Feinberg, E.; Jayatunga, B.H.D.; Lambrecht, W.R.L.; Zhu, M.; Hwang, J.; Kash, K.; et al. Deep level defects and cation sublattice disorder in ZnGeN₂. *J. Appl. Phys.* **2020**, *127*, 135703. [[CrossRef](#)]
16. Novikov, V.V.; Matovnikov, A.V.; Mitroshenkov, N.V.; Shevelkov, A.V.; Bud'ko, S.L. Crystal lattice disorder and characteristic features of the low-temperature thermal properties of higher borides. *Dalton Trans.* **2020**, *49*, 2138–2144. [[CrossRef](#)] [[PubMed](#)]
17. Marlton, F.P.; Zhang, Z.; Zhang, Y.; Proffen, T.E.; Ling, C.D.; Kennedy, B.J. Lattice Disorder and Oxygen Migration Pathways in Pyrochlore and Defect-Fluorite Oxides. *Chem. Mater.* **2021**, *33*, 1407–1415. [[CrossRef](#)]
18. Zhang, X.; Wang, H.; Hickel, T.; Rogal, J.; Li, Y.; Neugebauer, J. Mechanism of collective interstitial ordering in Fe–C alloys. *Nat. Mater.* **2020**, *19*, 849–854. [[CrossRef](#)]
19. Zhang, Y.Y.; Mishra, R.; Pennycook, T.J.; Borisevich, A.Y.; Pennycook, S.J.; Pantelides, S.T. Oxygen disorder, a way to accommodate large epitaxial strains in oxides. *Adv. Mater. Interfaces* **2015**, *3*, 1500344. [[CrossRef](#)]
20. Wang, X.; Ling, Y.; Lian, X.K.; Xin, Y.; Dhungana, K.B.; Perez-Orive, F.; Knox, J.M.; Chen, Z.; Zhou, Y.; Beery, D.; et al. Suppressed phase separation of mixed-halide perovskites confined in endotaxial matrices. *Nat. Commun.* **2019**, *10*, 695. [[CrossRef](#)]
21. Švecová, L.; Tillová, E.; Kuchariková, L.; Knap, V. Possibilities of predicting undesirable iron intermetallic phases in secondary Al-alloys. *Transp. Res. Procedia* **2021**, *55*, 797–804. [[CrossRef](#)]
22. Sopiha, K.V.; Larsen, J.K.; Donzel-Gargand, O.; Khavari, F.; Keller, J.; Edoff, M.; Platzer-Björkman, C.; Persson, C.; Scragg, J.J.S. Thermodynamic stability, phase separation and Ag grading in (Ag,Cu)(In,Ga)Se₂ solar absorbers. *J. Mater. Chem. A* **2020**, *8*, 8740–8751. [[CrossRef](#)]
23. Berche, A.; Tédenac, J.C.; Jund, P. Phase separation in the half-Heusler thermoelectric materials (Hf,Ti,Zr)NiSn. *Scr. Mater.* **2017**, *139*, 122–125. [[CrossRef](#)]
24. Synoradzki, K.; Ciesielski, K.; Kepiński, L.; Kaczorowski, D. Effect of secondary LuNiSn phase on thermoelectric properties of half-Heusler alloy LuNiSb. *Mater. Today Proc.* **2019**, *8*, 567–572. [[CrossRef](#)]
25. Gürth, M.; Rogl, G.; Romaka, V.; Grytsiv, A.; Bauer, E.; Rogl, P. Thermoelectric high ZT half-Heusler alloys Ti_{1-x-y}Zr_xHf_yNiSn (0 ≤ x ≤ 1; 0 ≤ y ≤ 1). *Acta Mater.* **2016**, *104*, 210–222. [[CrossRef](#)]
26. Dalpian, G.M.; Zhao, X.G.; Kazmerski, L.; Zunger, A. Formation and Composition-Dependent Properties of Alloys of Cubic Halide Perovskites. *Chem. Mater.* **2019**, *31*, 2497–2506. [[CrossRef](#)]
27. Chen, L.; Tan, Y.Y.; Chen, Z.X.; Wang, T.; Hu, S.; Nan, Z.A.; Xie, L.Q.; Hui, Y.; Huang, J.X.; Zhan, C.; et al. Toward Long-Term Stability: Single-Crystal Alloys of Cesium-Containing Mixed Cation and Mixed Halide Perovskite. *J. Am. Chem. Soc.* **2019**, *141*, 1665–1671. [[CrossRef](#)] [[PubMed](#)]
28. Sato, M.; Chai, Y.W.; Kimura, Y. Effect of Half-Heusler Interfacial Structure on Thermal Transport Properties of (Ti, Zr)NiSn Alloys. *ACS Appl. Mater. Interfaces* **2021**, *13*, 25503–25512. [[CrossRef](#)] [[PubMed](#)]
29. Bae, K.W.; Hwang, J.Y.; Kim, S.i.; Jeong, H.M.; Kim, S.; Lim, J.H.; Kim, H.S.; Lee, K.H. Thermoelectric Transport Properties of n-Type Sb-doped (Hf,Zr,Ti)NiSn Half-Heusler Alloys Prepared by Temperature-Regulated Melt Spinning and Spark Plasma Sintering. *Appl. Sci.* **2020**, *10*, 4963. [[CrossRef](#)]
30. El-Khouly, A.; Novitskii, A.; Adam, A.; Sedegov, A.; Kalugina, A.; Pankratova, D.; Karpenkov, D.; Khovaylo, V. Transport and thermoelectric properties of Hf-doped FeVsb half-Heusler alloys. *J. Alloys Compd.* **2020**, *820*, 153413. [[CrossRef](#)]
31. Bos, J.W.; Downie, R. Half-Heusler thermoelectrics: A complex class of materials. *J. Condens. Matter Phys.* **2014**, *26*, 433201. [[CrossRef](#)] [[PubMed](#)]
32. Luo, T.; Serrano-Sánchez, F.; Bishara, H.; Zhang, S.; Bueno Villoro, R.; Kuo, J.J.; Felsler, C.; Scheu, C.; Snyder, G.J.; Best, J.P.; et al. Dopant-segregation to grain boundaries controls electrical conductivity of n-type NbCo(Pt)Sn half-Heusler alloy mediating thermoelectric performance. *Acta Mater.* **2021**, *217*, 117147. [[CrossRef](#)]
33. Fu, C.; Bai, S.; Liu, Y.; Tang, Y.; Chen, L.; Zhao, X.; Zhu, T. Realizing high figure of merit in heavy-band p-type half-Heusler thermoelectric materials. *Nat. Commun.* **2015**, *6*, 8144. [[CrossRef](#)] [[PubMed](#)]
34. He, R.; Kraemer, D.; Mao, J.; Zeng, L.; Jie, Q.; Lan, Y.; Li, C.; Shuai, J.; Kim, H.S.; Liu, Y.; et al. Achieving high power factor and output power density in p-type half-Heuslers Nb_{1-x}Ti_xFeSb. *PNAS* **2016**, *113*, 13576–13581. [[CrossRef](#)]
35. Katre, A.; Carrete, J.; Mingo, N. Unraveling the dominant phonon scattering mechanism in thermoelectric compound ZrNiSn. *J. Mater. Chem. A* **2016**, *4*, 15940–15944. [[CrossRef](#)]
36. Naydenov, G.; Hasnip, P.; Lazarov, V.; Probert, M. Huge power factor in p-type half-Heusler alloys NbFeSb and TaFeSb. *J. Phys. Mater.* **2019**, *2*, 035002. [[CrossRef](#)]
37. Khandy, S. Inspecting the electronic structure and thermoelectric power factor of novel p-type half-Heuslers. *Sci. Rep.* **2021**, *11*, 1. [[CrossRef](#)]
38. Carrete Montaña, J.; Mingo, N.; Wang, S.; Curtarolo, S. Nanograined half-Heusler semiconductors as Advanced Thermoelectrics: An Ab Initio High-Throughput Statistical Study. *Adv. Funct. Mater.* **2014**, *24*, 7427–7432. [[CrossRef](#)]
39. El-Khouly, A.; Adam, A.; Ibrahim, E.; Nafady, A.; Karpenkov, D.; Novitskii, A.; Voronin, A.; Khovaylo, V.; Elsehly, E. Mechanical and thermoelectric properties of FeVsb-based half-Heusler alloys. *J. Alloys Compd.* **2021**, *886*, 161308. [[CrossRef](#)]
40. Ning, S.; Huang, S.; Zhang, Z.; Zhang, R.; Qi, N.; Chen, Z. High thermoelectric performance of topological half-Heusler compound LaPtBi achieved by hydrostatic pressure. *Phys. Chem. Chem. Phys.* **2020**, *22*, 14621–14629. [[CrossRef](#)]
41. Shiomi, J.; Esfarjani, K.; Chen, G. Thermal conductivity of half-Heusler compounds from first-principles calculations. *Phys. Rev. B* **2011**, *84*, 104302. [[CrossRef](#)]

42. Hoat, D. Electronic structure and thermoelectric properties of Ta-based half-Heusler compounds with 18 valence electrons. *Comput. Mater. Sci.* **2019**, *159*, 470–477. [[CrossRef](#)]
43. Fava, M.; Dongre, B.; Carrete, J.; van Roekeghem, A.; Madsen, G.K.H.; Mingo, N. Effects of doping substitutions on the thermal conductivity of half-Heusler compounds. *Phys. Rev. B* **2021**, *103*, 174112. [[CrossRef](#)]
44. Sun, H.L.; Yang, C.L.; Wang, M.S.; Ma, X.G. Remarkably High Thermoelectric Efficiencies of the Half-Heusler Compounds BXGa (X = Be, Mg, and Ca). *ACS Appl. Mater. Interfaces* **2020**, *12*, 5838–5846. [[CrossRef](#)]
45. Gaultois, M.W.; Sparks, T.D. How much improvement in thermoelectric performance can come from reducing thermal conductivity? *Appl. Phys. Lett.* **2014**, *104*, 113906. [[CrossRef](#)]
46. Wolf, M.; Hinterding, R.; Feldhoff, A. High power factor vs. high zT —A review of thermoelectric materials for high-temperature application. *Entropy* **2019**, *21*, 1058. [[CrossRef](#)]
47. Mao, J.; Wang, Y.; Kim, H.S.; Liu, Z.; Saparamadu, U.; Tian, F.; Dahal, K.; Sun, J.; Chen, S.; Liu, W.; et al. High thermoelectric power factor in Cu–Ni alloy originate from potential barrier scattering of twin boundaries. *Nano Energy* **2015**, *17*, 279–289. [[CrossRef](#)]
48. Skoug, E.; Zhou, C.; Pei, Y.; Morelli, D.T. High thermoelectric power factor in alloys based on CoSi. *Appl. Phys. Lett.* **2009**, *94*, 022115. [[CrossRef](#)]
49. Heinrich, C.; Day, T.; Zeier, W.; Snyder, G.; Tremel, W. Effect of isovalent substitution on the thermoelectric properties of the $\text{Cu}_2\text{ZnGeSe}_{4-x}\text{S}_x$ series of solid solutions. *J. Am. Chem. Soc.* **2013**, *136*, 442–448. [[CrossRef](#)] [[PubMed](#)]
50. Sun, H.; Lu, X.; Morelli, D.T. Isovalent substitutes play in different ways: Effects of isovalent substitution on the thermoelectric properties of $\text{CoSi}_{0.98}\text{B}_{0.02}$. *J. Appl. Phys.* **2016**, *120*, 035107. [[CrossRef](#)]
51. Chmielowski, S.; Bhattacharya, S.; Jacob, S.; Péré, D.; Jacob, A.; Moriya, K.; Delatouche, B.; Roussel, P.; Madsen, G.; Dennler, G. Strong reduction of thermal conductivity and enhanced thermoelectric properties in $\text{CoSb}_{1-x}\text{Se}_x$ paracostibite. *Sci. Rep.* **2017**, *7*, 46630. [[CrossRef](#)]
52. Shuai, J.; Wang, Y.; Liu, Z.; Kim, H.S.; Mao, J.; Sui, J.; Ren, Z. Enhancement of thermoelectric performance of phase pure Zintl compounds $\text{Ca}_{1-x}\text{Yb}_x\text{Zn}_2\text{Sb}_2$, $\text{Ca}_{1-x}\text{Eu}_x\text{Zn}_2\text{Sb}_2$, and $\text{Eu}_{1-x}\text{Yb}_x\text{Zn}_2\text{Sb}_2$ by mechanical alloying and hot pressing. *Nano Energy* **2016**, *25*, 136–144. [[CrossRef](#)]
53. Zevalkink, A.; Pomrehn, G.S.; Johnson, S.; Swallow, J.; Gibbs, Z.M.; Snyder, G.J. Influence of the triel elements (M = Al, Ga, In) on the transport properties of $\text{Ca}_5\text{M}_2\text{Sb}_6$ Zintl compounds. *Chem. Mater.* **2012**, *24*, 2091–2098. [[CrossRef](#)]
54. Guo, M.; Guo, F.; Zhu, J.; Yin, L.; Zhang, Q.; Cai, W.; Sui, J. Achieving high thermoelectric performance in rare-earth element-free CaMg_2Bi_2 with high carrier mobility and ultralow lattice thermal conductivity. *Research* **2020**, *2020*, 1–10. [[CrossRef](#)]
55. Zhou, J.; Huang, L.; Wang, Z.; Liu, Z.; Cai, W.; Sui, J. Effect of Cd isoelectronic substitution on thermoelectric properties of $\text{Zn}_{0.995}\text{Na}_{0.005}\text{Sb}$. *J. Mater.* **2016**, *2*, 324–330. [[CrossRef](#)]
56. Zheng, W.; Lu, Y.; Li, Y.; Wang, J.; Hou, Z.; Shao, X. Structural and thermoelectric properties of Zr-doped TiPdSn half-Heusler compound by first-principles calculations. *Chem. Phys. Lett.* **2020**, *741*, 137055. [[CrossRef](#)]
57. Zhang, J.; Huang, L.; Zhu, C.; Zhou, C.; Jabar, B.; Li, J.; Zhu, X.; Wang, L.; Song, C.; Xin, S.; et al. Design of domain structure and realization of ultralow thermal conductivity for record-high thermoelectric performance in chalcopyrite. *Adv. Mater.* **2019**, *31*, 1905210. [[CrossRef](#)]
58. Yu, J.; Fu, C.; Liu, Y.; Xia, K.; Aydemir, U.; Chasapis, T.; Snyder, G.; Zhao, X.; Zhu, T. Unique role of refractory Ta alloying in enhancing the figure of merit of NbFeSb thermoelectric materials. *Adv. Energy Mater.* **2017**, *8*, 1701313. [[CrossRef](#)]
59. Chauhan, N.S.; Bathula, S.; Vishwakarma, A.; Bhardwaj, R.; Johari, K.K.; Gahtori, B.; Saravanan, M.; Dhar, A. Compositional tuning of ZrNiSn half-Heusler alloys: Thermoelectric characteristics and performance analysis. *J. Phys. Chem.* **2018**, *123*, 105–112. [[CrossRef](#)]
60. Mallick, M.; Rajput, K.; Vitta, S. Increasing figure-of-merit of ZrNiSn half-Heusler alloy by minimal substitution and thermal conductivity reduction. *J. Mater. Sci. Mater.* **2019**, *30*, 6139–6147. [[CrossRef](#)]
61. Eliassen, S.N.H.; Katre, A.; Madsen, G.K.H.; Persson, C.; Løvvik, O.M.; Berland, K. Lattice thermal conductivity of $\text{Ti}_x\text{Zr}_y\text{Hf}_{1-x-y}\text{NiSn}$ half-Heusler alloys calculated from first principles: Key role of nature of phonon modes. *Phys. Rev. B* **2017**, *95*, 045202. [[CrossRef](#)]
62. Schrade, M.; Berland, K.; Eliassen, S.; Guzik, M.; Echevarria-Bonet, C.; Sørby, M.; Jenus, P.; Hauback, B.; Tofan, R.; Gunnæs, A.; et al. The role of grain boundary scattering in reducing the thermal conductivity of polycrystalline XNiSn (X = Hf, Zr, Ti) half-Heusler alloys. *Sci. Rep.* **2017**, *7*, 1. [[CrossRef](#)]
63. Yan, R.; Xie, W.; Balke, B.; Chen, G.; Weidenkaff, A. Realizing p-type NbCoSn half-Heusler compounds with enhanced thermoelectric performance via Sc substitution. *Sci. Technol. Adv. Mater.* **2020**, *21*, 122–130. [[CrossRef](#)] [[PubMed](#)]
64. Ono, Y.; Inayama, S.; Adachi, H.; Kajitani, T. Thermoelectric properties of NbCoSn -based half-Heusler alloys. In Proceedings of the 2006 25th International Conference on Thermoelectrics, Vienna, Austria, 6–10 August 2006; pp. 124–127. [[CrossRef](#)]
65. Kumar, A.; Chaturvedi, K.M.; Bano, S.; Govind, B.; Misra, D. Enhanced thermoelectric performance of p-type $\text{ZrCoSb}_{0.9}\text{Sn}_{0.1}$ via Tellurium doping. *Mater. Chem. Phys.* **2021**, *258*, 123915. [[CrossRef](#)]
66. Yuan, B.; Wang, B.; Huang, L.; Lei, X.; Zhao, L.; Wang, C.; Zhang, Q. Effects of Sb Substitution by Sn on the Thermoelectric Properties of ZrCoSb . *J. Electron. Mater.* **2016**, *46*. [[CrossRef](#)]
67. Tamura, S.I. Isotope scattering of dispersive phonons in Ge. *Phys. Rev. B* **1983**, *27*, 858–866. [[CrossRef](#)]

68. Hellman, O.; Abrikosov, I.A. Temperature-dependent effective third-order interatomic force constants from first principles. *Phys. Rev. B* **2013**, *88*, 144301. [[CrossRef](#)]
69. Hellman, O.; Abrikosov, I.A.; Simak, S.I. Lattice dynamics of anharmonic solids from first principles. *Phys. Rev. B* **2011**, *84*, 180301. [[CrossRef](#)]
70. Srivastava, G. *The Physics of Phonons*; Taylor & Francis: Abingdon, UK, 1990.
71. Feng, Z.; Fu, Y.; Zhang, Y.; Singh, D.J. Characterization of rattling in relation to thermal conductivity: Ordered half-Heusler semiconductors. *Phys. Rev. B* **2020**, *101*, 064301. [[CrossRef](#)]
72. Tranås, R.; Løvvik, O.M.; Tomic, O.; Berland, K. Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis. *Comput. Mater. Sci.* **2022**, *202*, 110938. [[CrossRef](#)]
73. Kresse, G.; Hafner, J. *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B* **1993**, *47*, 558–561. [[CrossRef](#)] [[PubMed](#)]
74. Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50. [[CrossRef](#)]
75. Kresse, G.; Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186. [[CrossRef](#)] [[PubMed](#)]
76. Perdew, J.P.; Ruzsinszky, A.; Csonka, G.I.; Vydrov, O.A.; Scuseria, G.E.; Constantin, L.A.; Zhou, X.; Burke, K. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406. [[CrossRef](#)]
77. Csonka, G.I.; Perdew, J.P.; Ruzsinszky, A.; Philipsen, P.H.T.; Lebègue, S.; Paier, J.; Vydrov, O.A.; Ángyán, J.G. Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **2009**, *79*, 155107. [[CrossRef](#)]
78. Shulumba, N.; Hellman, O.; Minnich, A.J. Intrinsic localized mode and low thermal conductivity of PbSe. *Phys. Rev. B* **2017**, *95*, 014302. [[CrossRef](#)]
79. Anderson, O. A simplified method for calculating the debye temperature from elastic constants. *J. Phys. Chem. Solids* **1963**, *24*, 909–917. [[CrossRef](#)]
80. Berland, K.; Shulumba, N.; Hellman, O.; Persson, C.; Løvvik, O.M. Thermoelectric transport trends in group 4 half-Heusler alloys. *J. Appl. Phys.* **2019**, *126*, 145102. [[CrossRef](#)]
81. Goldsmid, H.; Penn, A. Boundary scattering of phonons in solid solutions. *Phys. Lett. A* **1968**, *27*, 523–524. [[CrossRef](#)]
82. Tian, Z.; Lee, S.; Chen, G. Heat Transfer in Thermoelectric Materials and Devices. *J. Heat Transfer* **2013**, *135*, 061605. [[CrossRef](#)]
83. Xu, Q.; Zhou, J.; Liu, T.H.; Chen, G. Effect of electron-phonon interaction on lattice thermal conductivity of SiGe alloys. *Appl. Phys. Lett.* **2019**, *115*, 023903. [[CrossRef](#)]
84. Arrigoni, M.; Carrete, J.; Mingo, N.; Madsen, G.K.H. First-principles quantitative prediction of the lattice thermal conductivity in random semiconductor alloys: The role of force-constant disorder. *Phys. Rev. B* **2018**, *98*, 115205. [[CrossRef](#)]
85. Geng, H.; Zhang, H. Effects of phase separation on the thermoelectric properties of (Ti, Zr, Hf)NiSn half-Heusler alloys. *J. Appl. Phys.* **2014**, *116*, 033708. [[CrossRef](#)]
86. Rausch, E.; Balke, B.; Stahlhofen, J.M.; Ouardi, S.; Burkhardt, U.; Felser, C. Fine tuning of thermoelectric performance in phase-separated half-Heusler compounds. *J. Mater. Chem. C* **2015**, *3*, 10409–10414. [[CrossRef](#)]
87. Serrano-Sánchez, F.; Luo, T.; Yu, J.; Xie, W.; Le, C.; Auffermann, G.; Weidenkaff, A.; Zhu, T.; Zhao, X.; Alonso, J.A.; et al. Thermoelectric properties of n-type half-Heusler NbCoSn with heavy-element Pt substitution. *J. Mater. Chem. A* **2020**, *8*, 14822–14828. [[CrossRef](#)]
88. Chauhan, N.S.; Bathula, S.; Vishwakarma, A.; Bhardwaj, R.; Johari, K.K.; Gahtori, B.; Dhar, A. Enhanced thermoelectric performance in p-type ZrCoSb based half-Heusler alloys employing nanostructuring and compositional modulation. *J. Mater.* **2019**, *5*, 94–102. [[CrossRef](#)]
89. Hazama, H.; Matsubara, M.; Asahi, R.; Takeuchi, T. Improvement of thermoelectric properties for half-Heusler TiNiSn by interstitial Ni defects. *J. Appl. Phys.* **2011**, *110*, 063710. [[CrossRef](#)]
90. Xie, H.; Wang, H.; Fu, C.; Liu, Y.; Snyder, G.; Zhao, X.; Zhu, T. The intrinsic disorder related alloy scattering in ZrNiSn half-Heusler thermoelectric materials. *Sci. Rep.* **2014**, *4*, 6888. [[CrossRef](#)]
91. Ferluccio, D.A.; Kennedy, B.F.; Barczak, S.A.; Popuri, S.R.; Murray, C.; Pollet, M.; Bos, J.W.G. Thermal properties of TiNiSn and VFeSb half-Heusler thermoelectrics from synchrotron x-ray powder diffraction. *J. Phys. Energy* **2021**, *3*, 035001. [[CrossRef](#)]
92. Fu, C.; Xie, H.; Zhu, T.J.; Xie, J.; Zhao, X.B. Enhanced phonon scattering by mass and strain field fluctuations in Nb substituted FeVSb half-Heusler thermoelectric materials. *J. Appl. Phys.* **2012**, *112*, 124915. [[CrossRef](#)]
93. Cho, J.; Park, T.; Bae, K.; Kim, H.S.; Choi, S.M.; Kim, S.I.; Kim, S. Ti Addition Effect on the Grain Structure Evolution and Thermoelectric Transport Properties of Hf_{0.5}Zr_{0.5}NiSn_{0.98}Sb_{0.02} Half-Heusler Alloy. *Materials* **2021**, *14*, 4029. [[CrossRef](#)] [[PubMed](#)]
94. Zhao, H.; Cao, B.; Li, S.; Liu, N.; Shen, J.; Li, S.; Jian, J.; Gu, L.; Pei, Y.; Snyder, G.; et al. Engineering the Thermoelectric Transport in Half-Heusler Materials through a Bottom-Up Nanostructure Synthesis. *Adv. Energy Mater.* **2017**, *7*, 1700446. [[CrossRef](#)]
95. Su, C.H. Experimental determination of lattice thermal conductivity and Lorenz number as functions of temperature for n-type PbTe. *Mater. Today Phys.* **2018**, *5*, 58–63. [[CrossRef](#)]
96. Lu, Y.; Sun, T.; Zhang, D.B. Lattice anharmonicity, phonon dispersion, and thermal conductivity of PbTe studied by the phonon quasiparticle approach. *Phys. Rev. B* **2018**, *97*, 174304. [[CrossRef](#)]
97. Xiao, Y.; Chang, C.; Pei, Y.; Wu, D.; Peng, K.; Zhou, X.; Gong, S.; He, J.; Zhang, Y.; Zeng, Z.; et al. Origin of low thermal conductivity in SnSe. *Phys. Rev. B* **2016**, *94*, 125203. [[CrossRef](#)]

98. Yan, R.; Xie, R.; Xie, W.; Shen, C.; Li, W.; Balke, B.; Yoon, S.; Zhang, H.; Weidenkaff, A. Effects of Doping Ni on the Microstructures and Thermoelectric Properties of Co-Excessive NbCoSn Half-Heusler Compounds. *ACS Appl. Mater. Interfaces* **2021**, *13*, 34533–34542. [[CrossRef](#)]
99. Dai, C.K.; Song, Q.F.; Xie, L.; Liu, R.H.; Bai, S.; Chen, L.D. Improving thermoelectric properties of ZrPtSn-based half-Heusler compound by Sb doping. *Rare Met.* **2021**, *40*, 2838–2846. [[CrossRef](#)]
100. Lee, S.; Esfarjani, K.; Luo, T.; Zhou, J.; Tian, Z.; Chen, G. Resonant bonding leads to low lattice thermal conductivity. *Nat. Commun.* **2014**, *5*, 3525. [[CrossRef](#)] [[PubMed](#)]

Paper III

Lattice thermal conductivity of cubic compounds with iterative training over material space using Gaussian process regression

R. Tranås, O. M. Løvvik, and K. Berland

TO BE SUBMITTED



Lattice thermal conductivity of cubic compounds with iterative training over material space using Gaussian process regression

Rasmus Tranås,¹ Ole Martin Løvvik,^{2,3} and Kristian Berland¹

¹*Department of Mechanical Engineering and Technology Management, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

²*Centre for Materials Science and Nanotechnology,*

Department of Physics, University of Oslo, NO-0316 Oslo, Norway

³*SINTEF Sustainable Energy Technology, NO-0314 Oslo, Norway*

(Dated: December 19, 2022)

The lattice thermal conductivity is a key material property in power electronics, thermal barriers, and thermoelectric devices. Identifying a wide range of low lattice thermal conductivity compounds is important for thermoelectric power generation, due to the many criteria for such materials. Here, we present a machine learning model using an active learning scheme based on Gaussian process regression, which is used to accurately predict the lattice thermal conductivity of cubic compounds. The features are based on atomic and structural information to enable efficient screening of materials databases and hypothetical compounds. The lattice thermal conductivity used for training is calculated using density functional theory and the temperature-dependent effective potential method. Iteratively adding 30 compounds based on the uncertainty measure of the Gaussian process regression to the training set reduces prediction uncertainty up to $\sim 35\%$. The machine learning model is used to screen 1573 compounds in the MATERIALS PROJECT database. Among the compounds with band gaps, 34 have predicted lattice thermal conductivity ≤ 1.3 W/Km at 300 K. Two materials identified, Na_2TlSb and Ca_3AsBr_3 , are studied to obtain insight into the origin of the low lattice thermal conductivity. Our study highlights the benefit of data-driven active sampling. The machine learning model and active sampling scheme can be extended to broader classes of systems in a systematic exploration of low lattice thermal conductivity compounds.

I. INTRODUCTION

The thermal conductivity of semiconductors and technical ceramics is a key material parameter for many applications. For power electronics, high thermal conductivity is needed to divert heat away from transistor components, while thermal barrier materials in jet engines or thermoelectrics, require low thermal conductivity. Thermoelectric devices, with their ability to convert heat into electricity, could play a role in reducing global energy consumption by harvesting waste heat [1–3].

The efficiency of thermoelectric materials is given by the figure-of-merit, ZT , which at temperature T is given as

$$ZT = \frac{\sigma S^2 T}{\kappa_\ell + \kappa_e}, \quad (1)$$

where σ is the electrical conductivity, S is the Seebeck coefficient, κ_e is the electronic thermal conductivity, and κ_ℓ is the lattice thermal conductivity. A high ZT material must both have a low value of κ_ℓ and good electronic properties. The latter are highly interlinked; increasing the value of σ by increasing doping concentration, reduces S and also increases κ_e [4]. The value of κ_ℓ and the electronic properties are less connected, but not entirely so. For instance, the half-Heuslers [5], can have favorable electronic properties but suffer from larger intrinsic lattice thermal conductivity. These compounds can be competitive by taking measures to reduce lattice thermal conductivity, such as by increasing phonon

scattering from introducing grain boundaries [6, 7], point defects [8, 9], or alloying elements [4, 10–13]. Still, such measures tend to also decrease electronic mobility and thus also σ [14–16]. Materials with intrinsically low κ_ℓ are therefore attractive. High performance thermoelectric compounds, such as PbTe [17, 18], Bi_2Te_3 [19, 20], and SnSe [21, 22], can attain $\kappa_\ell < 1$ W/Km at room temperature. However, materials with both low κ_ℓ and attractive electronic properties are quite rare. It is therefore important to map out many material candidates with low κ_ℓ . In addition to the many competing criteria for high performance, it is desirable to identify environmentally friendly and cheap materials that can be realized with conventional synthesis methods. Many of the most widely used thermoelectric materials contain toxic or rare earth elements [23, 24].

Obtaining experimental lattice thermal conductivity is a time-consuming task requiring expert skills and costly equipment, reducing the number of compounds studied. In the literature, lattice thermal conductivity for ~ 100 –200 compounds are reported [5, 25, 26]. High-throughput computational screening using density functional theory (DFT) has in recent years accelerated the identification of low κ_ℓ compounds [27–31] and there is a growing number of computed κ_ℓ values in the literature [32, 33]. Nonetheless, calculation of κ_ℓ requires substantial computer resources. The large cost comes from the high number of large supercell calculations with DFT needed for computing the three-phonon scattering rates [34, 35]. Recent works have used techniques such as compressive sensing [36] or machine learning (ML) [37] to reduce the

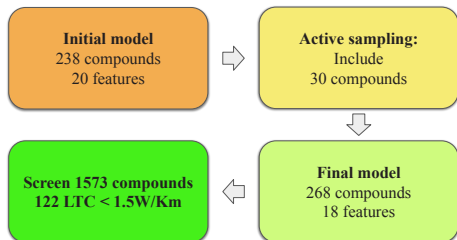


FIG. 1: Flowchart for model construction and screening.

computational cost by limiting the number of configurations required to be treated with DFT. ML is increasingly also being adopted to predict physical properties directly in thermal energy research [38] and for screening of low lattice thermal conductivity compounds [32, 39–47]. Such methods became of particular interest after the advent of structured material databases, such as MATERIALS PROJECT [48], OQMD [49], and NOMAD [50], as they enable effective screening on compounds that are readily formatted for use with ML.

ML training sets for lattice thermal conductivity are typically quite modest in size, i.e. covering ~ 40 -200 compounds [25, 26, 32, 43], but it is often applied to 1000s or even 10000s of compounds. This low ratio is an issue because it becomes highly likely that some unseen compounds have significantly different properties than those in the training set [51]. Ensuring a training set representative of the set of unseen compounds is thus crucial for the accurate predictions needed to identify new low lattice thermal conductivity compounds [52]. The refinement of training sets i.e. active sampling – has gathered significant attention recently [53], such as in training ML force fields [54–57], due to the large computational cost of ab initio calculations. Additionally, when starting from literature data, which is common in ML, the training set easily becomes unbalanced. Certain material classes could be quite well sampled – such as the half-Heuslers, a starting point for this paper – while other material classes are more sparsely sampled.

In this work, we couple ML and active sampling to predict lattice thermal conductivity of 1573 cubic compounds in the MATERIALS PROJECT database. Fig. 1 outlines the workflow. Explicit lattice thermal conductivity calculations based on density functional theory (DFT) and the temperature-dependent effective potential (TDEP) method [35, 58] provide data for training, which together with Gaussian process regression enable active sampling, which results in an accurate ML model.

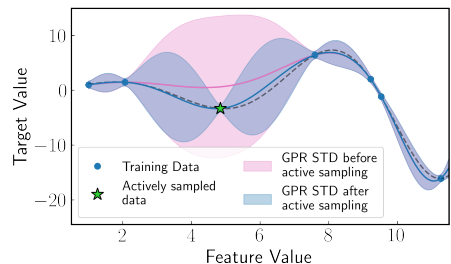


FIG. 2: One dimensional active sampling with GPR: feature axis on the horizontal and prediction on the vertical. The purple (blue) shaded region gives the GPR STD before (active) sampling.

II. METHODS

A. Gaussian Process Regression

Gaussian process regression (GPR) and classification have seen much usage in materials science [59, 60], including for training of force-fields [54, 61], for materials discovery [62], and for predicting thermal- [63] and lattice-thermal conductivity [25, 39]. A prediction with GPR, κ_ℓ^* , for a compound with features, \mathbf{x}_* , given a vector of training κ_ℓ , $\{\kappa_\ell\}$, and corresponding features, X , is based on a distribution $p(\kappa_\ell^* | \{\kappa_\ell\}, X, \mathbf{x}_*)$. The distribution provides GPR with a distribution mean and a variance, where the mean is the model prediction and the variance can be interpreted as the uncertainty of the prediction. Kernels incorporate the co-variance between compound features into the model and enable calculation of the distributions. When the features of two compounds are similar, $x \approx x'$, the co-variance is high and $k(x, x')$ is near its maximum. The GPR model used in this work is based on the Matern kernel, expressed as

$$k(x, x') = \frac{1}{\Gamma(\nu)2^{\nu-1}} (rd(x, x'))^\nu K_\nu(rd(x, x')). \quad (2)$$

Here, $\Gamma(\nu)$ is the gamma function, $d(x, x')$ is the Euclidean distance, K_ν is a modified Bessel function, and $r = \sqrt{2\nu}/\ell$, where ν is a free parameter. The length-scale parameter, ℓ , toggles how fast $k(x, x')$ drops off for increasing dissimilarity between compounds.

We employ an active sampling scheme based on uncertainties from GPR, the GPR standard deviation (STD) of predictions [64]. GPR STD-based active sampling has been used to determine structures that should be treated with ab initio methods instead of ML [55]. Fig. 2 shows GPR predictions and STD for a toy model trained on six data points (purple, before active sampling) and a model trained on seven data points (blue, after active sampling).

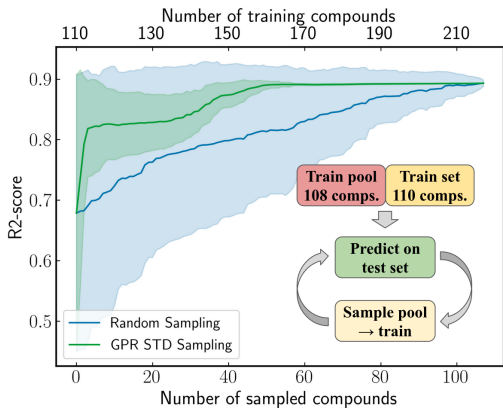


FIG. 3: R2-scores for predictions made on the test set (vertical axis) vs. number of compounds sampled to the training set (bottom horizontal axis). The top horizontal axis shows the number of training compounds.

When the new point to sample is added in the training the accuracy of the GPR predictions increases and the GPR STD reduces.

The active sampling scheme selects new compounds from the MATERIALS PROJECT iteratively. It starts by sampling the compound with the highest GPR STD and adding it to the training set with lattice thermal conductivity equal to the prediction from the GPR model. This process is repeated for 29 compounds. 15 of the 29 compounds have strictly positive phonon dispersions from TDEP at 300 K and are dynamically stable. For these compounds, the TDEP lattice thermal conductivity, $\kappa_{\ell}^{\text{TDEP}}$, is calculated, and the compounds are included in the training. This process of obtaining a block of 15 compounds is carried out once more in a similar fashion. In the second block, 15 out of 30 compounds are dynamically stable. The two blocks constitute 30 actively sampled compounds and their $\kappa_{\ell}^{\text{TDEP}}$.

B. Validation of Active sampling scheme

To validate the active sampling scheme, we construct three data sets based on the compounds presented in Sec. IID, a start training set, training pool, and test set (110, 108, and 20 compounds, respectively). In each iteration, a compound is moved from the training pool to the training set. For the random sampling model, the compound moved from pool to training is randomly selected from the pool, while for the active sampling model, the compound moved is the one in the training pool with the highest GPR STD. For each sampled compound, both models are retrained and used to predict on the test set.

Fig. 3 displays the R2-scores for predictions on the test set using active and random sampling. The blue line shows the R2-score of the model sampling random compounds to include in the training and the green line shows the score of the model sampling the compounds with the highest GPR STD. The shaded regions are standard deviations from 100 different random splits of the training set and training pool. At 0 and 108 on the horizontal axis, the training sets are identical (no compounds sampled, all compounds sampled), and the R2-scores for the models are the same. The active sampling scheme achieves a high R2-score with fewer samples than the random sampling while having a lower standard deviation.

C. Features and Dataset

Features used for ML are generated from material properties obtained directly from databases and the atomic structure. To account for the fact that different compounds have a different number of atoms, we define averages $\langle \cdot \rangle$, medians $\bar{m}(\cdot)$, and standard deviations $\sigma(\cdot)$, of different atomic properties, i.e. covalent radii [65], r , electronegativities [66], χ , masses [67], m , number of valence electrons, N_{val} , the number of s, p, d, and f electrons, N_s , N_p , N_d , and N_f in the valence shell [68], and dipole polarizabilities [69], μ_{dip} , where μ_{dip} describes the charge distribution response of the atom when an external electric field is applied. We also include the number of elements in the primitive cell, N_{ele} .

To represent information related to coordination and symmetry we construct features from the Voronoi structure [43, 70, 71] of the lattice. The Voronoi structure for an atom situated in a lattice is defined by planes perpendicular to the vector connecting the atom with its closest neighbors at the halfway point of the vector. Fig. 4 shows the Voronoi structure of a body-centered cubic lattice, with a perimeter highlighted in purple and a Voronoi polyhedron encapsulating the central atom.

As features, we use corresponding averages, medians, and standard deviation of the total length of all perimeter circumferences for the atom, $\sum_n L_{\text{per},n}$, the product of the perimeter circumference and the sum of the covalent radius of the atom and its neighbors, $\sum_n L_{\text{per},n}(r+r_n)$, the volume of the sphere defined by the covalent radius for the atom divided by the volume of the Voronoi polyhedron, $V_{\text{covrad}}/V_{\text{Vor}}$, the effective coordination number (ECN), $\text{ECN} = (\sum_n L_{\text{per},n})^2 / \sum_n L_{\text{per},n}^2$, the Voronoi weighted Steinhardt's parameters [72, 73], S_i , $i = (1, 2, 3, 4)$, and coordination number, N_{cor} . The Voronoi-based features are calculated using PYSICAL [74] and are scaled by the DFT calculated lattice parameter. The features are standardized by subtracting the mean and dividing by the standard deviation and summarized in Appendix A, Table II.

The features require no DFT calculations, enabling accelerated rapid screening of hypothetical materials databases such as the MATERIALS PROJECT [48]. Only

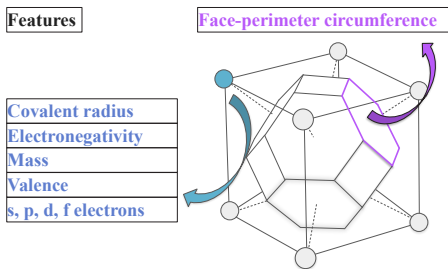


FIG. 4: Voronoi structure for the central atom and features obtained from atomic properties.

the atomic configuration and crystal structure are needed for feature construction. Higher-order features requiring calculations – e.g. volume, bulk modulus [25, 52], and properties related to anharmonicity [32, 75] – can improve ML performance, although the computational cost could limit the number of compounds studied. The MATERIALS PROJECT provides DFT computed volume for all compounds in the database and bulk modulus for a subset. We avoid using these features for two main reasons: The DFT parameters in the MATERIALS PROJECT vary from those used in this work, and using such features prevents employing the model on prototype compounds in future studies [43, 76, 77].

Functions used to construct features use the array of properties of individual atoms as input. Our chosen functions have good scalability to e.g. supercells with low-concentration impurities. Avoiding functions such as minimum and maximum reduces the likelihood that a feature changes abruptly in a supercell with impurities.

Including all available features in ML training can lead to overfitting and poor accuracy. Forward sequential feature floating selection (SFFS) [78] obtains a suitable feature subset by removing redundant features. This scheme iteratively includes the feature that maximizes the R2-score in the feature subset. When a feature is included all features in the subset are removed one-by-one and if the removal of a feature increases the R2-score it is removed. Model hyperparameters are tuned during each evaluation of a subset and 10-fold cross-validation (CV) is used. For the initial model, feature selection is done with SFFS using a set of 218 compounds, where 20 compounds are reserved as a test set. After completing the active sampling SFFS is done using 268 compounds. We use the code from MLXTEND [79] for the SFFS, while the GPR model is as implemented in SCIKIT-LEARN [80]. After feature selection, it is interesting to note the importance of individual features. In this work, we analyze the feature importance from feature permutation. The feature importance is defined as the reduction in R2-score when feature values are shuffled randomly, shuffling one feature at a time going through all features.

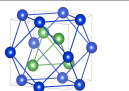
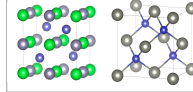




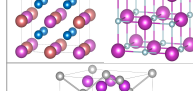
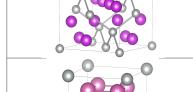

	Structure	Space group	Number of compounds
	gamma brass-derived	215	NS: 24 NI: 0 NA: 5
	half-Heusler, zinc blende	216	NS: 209 NI: 152 NA: 0
	tetrafluoride	217	NS: 2 NI: 0 NA: 0
	perovskite	221	NS: 358 NI: 32 NA: 5
	LaRuSn 3-type	223	NS: 4 NI: 0 NA: 2
	cuprite	224	NS: 4 NI: 0 NA: 2
	full-Heusler, rock salt	225	NS: 858 NI: 53 NA: 4
	Laves	227	NS: 103 NI: 1 NA: 7
	CsCl-derived	229	NS: 13 NI: 0 NA: 3

FIG. 5: Number of screened compounds (NS), compounds in the initial training set (NI), and actively sampled compounds (NA).

D. Compound Set

The initial compound set contains 238 cubic compounds. It consists of 122 half-Heusler compounds from Ref. [52] in addition to a random selection of 106 cubic compounds from the MATERIALS PROJECT [48]. We have also included known low lattice thermal conductivity compounds. These are CsI [81], CuCl [82], TiCl [83], TiBr [84], Cu₂S [85], Ag₂S [86], Ba₂BiAu [87, 88], Ba₂SbAu [89], Sr₂SbAu [89], and Ba₂HgPb [87].

The set of unknown compounds to be screened is a subset of the remaining cubic compounds in the MATERIALS PROJECT database [48, 90]. Fig. 5 summarizes the

compounds. We exclude compounds containing elements H, Yb, Lu, Po, and U. Compounds with more than eight atoms in the primitive cell, one atom in the primitive cell, or compounds with non-zero magnetic moments are also excluded. Lastly, compounds with energy higher than 0.1 eV/atom over the convex hull are removed from the data set.

E. Lattice Thermal Conductivity

The TDEP [35, 58] method enables extraction of effective second- and third-order force constants at finite temperature that are needed to compute lattice thermal conductivity. This method compares DFT forces with effective forces on configurations with atomic displacements to calculate force constants that best replicate DFT forces. From the phonon Boltzmann transport equation within the relaxation-time approximation lattice thermal conductivity, κ_ℓ , can be expressed as,

$$\kappa_\ell = \kappa_{\alpha\alpha} = \frac{1}{V} \sum_{\mathbf{q}s} c_{\mathbf{q}s} v_{\alpha\mathbf{q}s}^2 \tau_{\alpha\mathbf{q}s}. \quad (3)$$

Here, α is a Cartesian direction, V is the volume, \mathbf{q} is the wave vector, s is the branch, $c_{\mathbf{q}s}$ is the phonon heat capacity, $v_{\alpha\mathbf{q}s}$ is the group velocity, and $\tau_{\alpha\mathbf{q}s}$ is the lifetime. The $\kappa_\ell = \kappa_{\alpha\alpha}$ holds for cubic compounds as off-diagonal terms in the κ_ℓ tensor are zero and κ_ℓ in the three Cartesian directions is the same. DFT calculations are done with the VASP [91–93] software package. For compounds with four or less atoms in the primitive cell, we use the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation for solids, PBEsol [94, 95], while for the rest of the compounds we use the vdW-DF-cx functional [96, 97]. PBEsol enables computation of accurate lattice parameters and vibrational properties [98], and vdW-DF-cx provides accurate lattice parameters for cubic perovskites [99] and has good transferability between compounds [100, 101]. For relaxations, a 500 eV plane-wave energy cutoff and a $11 \times 11 \times 11$ \mathbf{k} -point (four or less atoms) or $9 \times 9 \times 9$ \mathbf{k} -point (five or more atoms) sampling is used. Self-consistency in the electronic iterations is reached at 10^{-6} eV, and atomic positions are relaxed until no forces are above 1 meV/Å. Second- and third-order force constants are extracted from DFT force calculations on fifty configurations of $3 \times 3 \times 3$ repetitions of the primitive cell ($4 \times 4 \times 4$ for compounds with two atoms in the primitive cell), using TDEP. These thermally excited configurations have displaced atomic positions corresponding to a canonical ensemble at $T = 300$ K [102]. The finite temperature displacements are obtained by matching the zero-point energy of the phonons to the Debye temperature calculated from the elastic tensor [103]. The cutoffs for the second- and third-order force constants are set to be $\sim 12\%$ and $\sim 2\%$ larger than half the width of the supercell. The \mathbf{q} -point grid for the phonon-mode integration in Eq. (3)

is discretized as $35 \times 35 \times 35$ for compounds with less than five atoms in the primitive cell and $30 \times 30 \times 30$ for the rest.

III. RESULTS AND DISCUSSION

A. Model Construction with Active Sample Selection

The initial GPR model is based on the 238 compounds presented in Sec. IID and their TDEP κ_ℓ , $\kappa_\ell^{\text{TDEP}}$. Feature selection with SFFS provides a feature subset that yields better model performance compared to using all features. Table II displays the 20 chosen features. The SFFS selects ten features based on the Voronoi tessellation structure, while the remaining are based on tabulated data for χ , N_{val} , r , μ_{dip} , m , N_{p} , and N_{ele} . The feature importance is also shown in Appendix A, Table II. The importance is higher than zero in all cases, showing that all features increase model performance. The top three features are the standard deviation of electronegativity, σ_χ , median number of valence electrons, $\bar{m}(N_{\text{val}})$, and the standard deviation of the product of perimeters and sum of covalent radii of neighbors, $\sigma(\sum_n L_{\text{per},n}(r+r_n))$. σ_χ correlates most with $\kappa_\ell^{\text{TDEP}}$, out of all features. The Spearman correlation is -0.46 , i.e. compounds with a large variation in χ of individual atoms tend to have lower $\kappa_\ell^{\text{TDEP}}$. The group 4 and 5 half-Heuslers have $\bar{m}(N_{\text{val}}) = 4$ or $\bar{m}(N_{\text{val}}) = 5$, as for TiNiSn and NbCoSb, which are differentiated from compounds such as zincblende AgBr and CuCl with $\bar{m}(N_{\text{val}}) = 9$. The median mass, $\bar{m}(m)$, is negatively correlated with $\kappa_\ell^{\text{TDEP}}$. A similar correlation has been seen when comparing the mean atomic mass and experimental lattice thermal conductivity [25]. The presence of heavy atoms can contribute to lowering the phonon group velocity and thus also κ_ℓ [104].

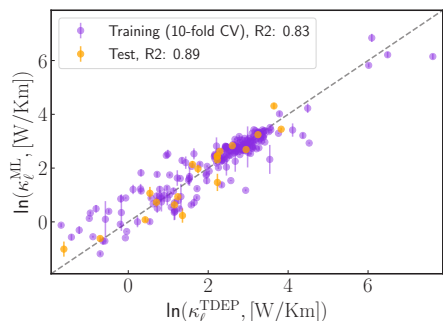


FIG. 6: Prediction accuracy for the ML model. The horizontal axis shows the logarithm of $\kappa_\ell^{\text{TDEP}}$ and the vertical axis shows the logarithm of κ_ℓ^{IML} .

Fig. 6 compares predictions from the validation of the initial ML model, κ_ℓ^{ML} , and $\kappa_\ell^{\text{TDEP}}$. The purple markers show the average predictions from a 10-fold CV repeated 10 times with shuffled training data. The yellow markers show the predictions of the test set, which is withheld during feature selection and hyperparameter optimization. The error bars show the respective standard deviations. The R2-score for the predictions on the training set is 0.83 and for the test set, 0.89. The Spearman correlation for the training set is 0.91 and for the test set, 0.95. The high correlations show that the model accurately ranks compounds from low to high $\kappa_\ell^{\text{TDEP}}$. In the following the model is trained on all 238 compounds with predictions κ_ℓ^{IML} .

Fig. 7 (a) shows the high-uncertainty compounds actively sampled from the set of cubic compounds from MATERIALS PROJECT. The GPR STD of the iteratively sampled compounds is shown as bars. Before sampling Hg₄Pt (first green bar from the left) the model is re-trained with the compounds corresponding to red bars and their $\kappa_\ell^{\text{TDEP}}$. The GPR STD increases between the two blocks from SnI₄ to Hg₄Pt, possibly resulting from the removal of unstable compounds and change in target values from ML predictions to $\kappa_\ell^{\text{TDEP}}$. The sampled compounds consist of 21 with more than five atoms in the primitive cell, 38 different elements, and 7 space groups. Fig. 5 shows that the active sampling yields compounds with space groups not present in the training set at the start of the sampling.

Fig. 7 (b) displays $\kappa_\ell^{\text{TDEP}}$ and κ_ℓ^{IML} for the sampled compounds. Appendix B Table III contains $\kappa_\ell^{\text{TDEP}}$ for the total 268 cubic compounds at 300 K. The data can also be found in the supplementary materials. I₄Sn, InTl₃, CsBi₂, have low $\kappa_\ell^{\text{TDEP}}$ of 0.41, 0.56, 0.79 W/Km, respectively. The highest $\kappa_\ell^{\text{TDEP}}$ is obtained for GeSi₇ with $\kappa_\ell^{\text{TDEP}} = 59.48$ W/Km. Cs₂Se has $\kappa_\ell^{\text{TDEP}} = 1.35$ W/Km and has been found in a previous screening study by Juneja et al. [41]. A comparison of $\kappa_\ell^{\text{TDEP}}$ and κ_ℓ^{IML} shows that κ_ℓ^{IML} is significantly higher than $\kappa_\ell^{\text{TDEP}}$ for I₄Sn and InTl₃, with relative errors of 270 and 500 %.

In Fig. 8 (a) we show the compounds to be screened in the space spanned by the normalized average mass, $\langle m \rangle$, and standard deviation of the electronegativity, σ_χ . The colorbar displays the GPR STD. In the left panel the initial model is used, in the middle panel the model is re-trained with $\kappa_\ell^{\text{TDEP}}$ from the first block of 15 actively sampled compounds (red stars), and in the right panel the model is re-trained with the full 30 compounds (red and green stars). Dark pockets in the left panel have high GPR STD, corresponding to high uncertainty compounds. The GPR STD decreases when compounds are sampled to the training set, seen as blue discs turn to gray/white discs. In the left panel, the average GPR STD of the 100 compounds with the highest GPR STD is 1.7, in the middle panel it is 1.2, and in the right panel, it is 1.1 – a 35 % reduction compared to the starting point. The large spread of the actively sampled compounds in Fig. 8 (a) and (b) shows that the active sampling yields

compounds with a large span of feature values.

A 10 times shuffled 10-fold CV including compounds from the active sampling with the previously obtained features gives an R2-score of 0.77 and Spearman correlation of 0.89. Feature selection with SFFS yields 18 features where 11 features are the same before and after active sampling. The features are shown in Table II. After SFFS the shuffled 10-fold CV yields predictions with an R2-score of 0.81 and Spearman correlation of 0.93.

Next, we compare predictions made by the initial model, κ_ℓ^{IML} , trained on 238 compounds, and the active sampling model, κ_ℓ^{AML} , trained on 268 compounds. There are 55 compounds with $\kappa_\ell^{\text{AML}} < 2$ W/Km that have $\kappa_\ell^{\text{IML}} > 4$ W/Km. Tl₃Bi and In₃Bi have $\kappa_\ell^{\text{AML}} \sim 80$ % lower than κ_ℓ^{IML} , with κ_ℓ^{AML} of 0.71 and 0.86 W/Km, respectively. Na₂TlSb has $\kappa_\ell^{\text{AML}} \sim 35$ % lower than κ_ℓ^{IML} , with $\kappa_\ell^{\text{AML}} = 0.6$ W/Km.

Fig. 9 shows $\kappa_\ell^{\text{TDEP}}$ for the training compounds and κ_ℓ^{AML} for the screened compounds. The range of $\kappa_\ell^{\text{TDEP}}$ covers the range of κ_ℓ^{AML} . κ_ℓ^{AML} spans three orders of magnitude, where the lowest is CsK₃ with $\kappa_\ell^{\text{AML}} = 0.26$ W/Km. BAs is a zincblende compound known for its high thermal conductivity [105] and is the compound with highest κ_ℓ^{AML} in the study with $\kappa_\ell^{\text{AML}} = 304.54$ W/Km. This compound has been researched for its potential in thermal management applications, and experimental thermal conductivity has been measured at 1300 W/Km [106]. Among the other compounds with κ_ℓ^{AML} on the higher end are also BeO with $\kappa_\ell^{\text{AML}} = 100.95$ W/Km, and BeSe with $\kappa_\ell^{\text{AML}} = 76.53$ W/Km. These two compounds have previously been reported with $\kappa_\ell = 250$ W/Km for BeO, and $\kappa_\ell = 633$ for BeSe, based on theoretical calculations. While κ_ℓ^{AML} significantly underestimates the previous values, the ML model can separate these high κ_ℓ compounds from the rest.

B. Screening of Low Lattice Thermal Conductivity Compounds

Table I lists the 34 out of the 1573 compounds which attain $\kappa_\ell^{\text{AML}} \leq 1.30$ W/Km and non-zero band gap, E_g , reported in the MATERIALS PROJECT. The supplementary materials contain κ_ℓ^{AML} for all compounds. An additional 89 compounds are listed with $E_g = 0$ eV in the MATERIALS PROJECT, which can make them less suitable for thermoelectric applications. Note that some of these might be incorrectly labeled as metallic due to usage of standard DFT in the generalized gradient approximation, rather higher level of theory, such as at the hybrid functional level. In fact, we recently identified Ba₂HgPb, which has very low lattice thermal conductivity (in the training set listed in Table III), as such a compound [107]. The E_g for the compounds in Table I ranges from 0.04 eV (Ba₂SnHg) to 4.39 eV (CsCaBr₃). The average mass of the compounds varies significantly, with the lightest having an average mass of 59.8 u (K₃Sb) and the heaviest having 165.6 u (TlI). The lowest ratio

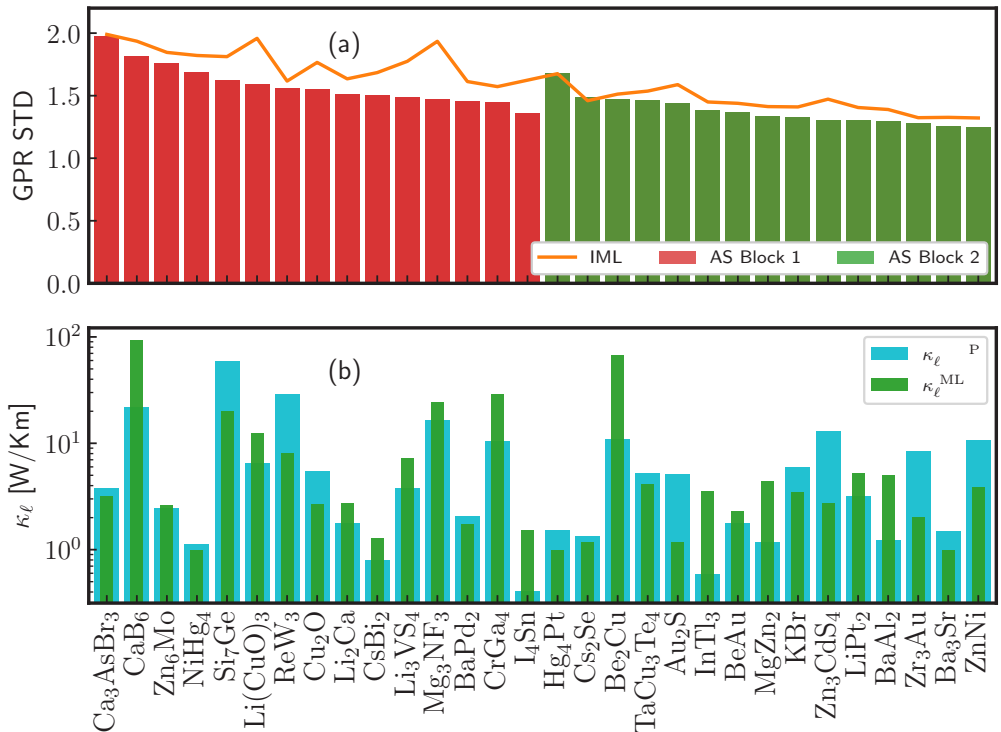


FIG. 7: (a) Bars show the GPR STD after each actively sampled compound is added to the training set. Red bars correspond the first active sampling (AS) block of 15 compounds, and green the second. The orange line shows the GPR STD from the initial model before active sampling. (b) κ_ℓ^{TDEP} and κ_ℓ^{IML} for the actively sampled compounds.

between the lightest and heaviest mass of elements in the compounds is 0.11 (Na_2TlBi) and the highest is 0.79 (Ba_2AgSb). The compounds range from 2 to 5 atoms in the primitive cell and span three different space groups: 221 (perovskite), 225 (full-Heusler), and 227 (Laves). Jaafreh et al. found several compounds containing Cs with low lattice thermal conductivity in their screening study [43]. The low lattice thermal conductivity was attributed to low phonon lifetimes and high anharmonicity. In the same fashion, in this work we identify 10 Cs-containing compounds to have $\kappa_\ell^{AML} < 1.30$ W/Km.

Seventeen of the compounds in Table I have not previously to our knowledge, been identified as low-lattice thermal conductivity compounds. These are shown with bold text, while seventeen other compounds have previously been researched. The full-Heusler Ba_2AgSb has $\kappa_\ell^{AML} = 0.47$ W/Km and κ_ℓ of 1.4 W/Km [108] has previously been calculated, where the low κ_ℓ was attributed to strong phonon anharmonicity. Ba_2AgSb and Ba_2AgBi have also shown to have low calculated κ_ℓ and

$ZT \approx 1$ [109]. He et al. identified a class of full-Heuslers X_2YZ , $X = Ca, Sr, \text{ and } Ba$, $Y = Au, Z = Sn, Pb, As, Sb, \text{ and } Bi$, with low κ_ℓ using high-throughput screening [87]. We find one compound with $X = Ca$, five with $X = Sr$, and four compounds with $X = Ba$ to have $\kappa_\ell^{AML} \leq 1.30$ W/Km.

The full-Heusler Na_2TiSb is one of the low lattice thermal conductivity compounds we identified, $\kappa_\ell^{AML} = 0.60$ W/Km. It is on the convex hull with $E_g = 0.38$ eV reported in MATERIALS PROJECT. A separate TDEP calculation is performed to validate the ML prediction, resulting in $\kappa_\ell^{TDEP} = 0.50$ W/Km. The standard deviation of the electronegativity of the atoms in the primitive cell is high, corresponding to a κ_ℓ^{TDEP} on the lower end. While previous research has identified low κ_ℓ full-Heuslers, Na_2TiSb distinguishes itself by having a very large mass ratio between the lightest and heaviest, i.e. Na and Tl elements.

Fig. 10 shows the phonon dispersions, phonon lifetimes, spectral κ_ℓ^{TDEP} , and site-projected density of

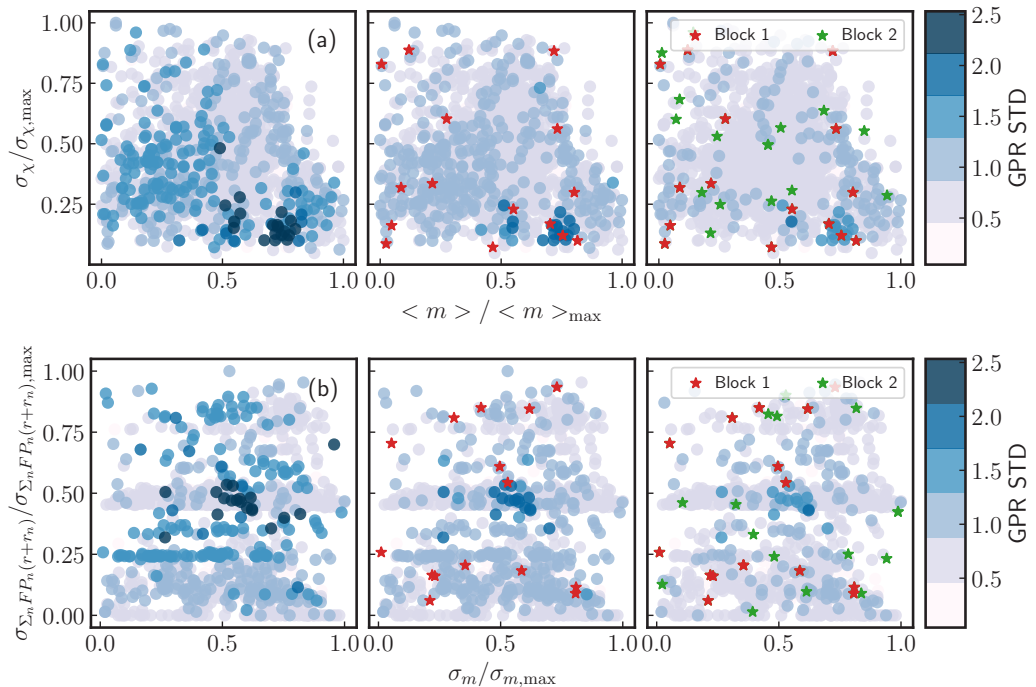


FIG. 8: (a) Evolution of uncertainties for compounds from MATERIALS PROJECT during active sampling. The horizontal and vertical axis show the normalized average mass and standard deviation of the electronegativity. The white-blue scale displays the GPR STD values. In the left panel, no additional compounds are sampled to the training set. In the middle panel, the first compound block is included in the training, indicated by red stars. For the right panel, the second block is also included, indicated by green stars. (b) shows the same as (a) with the normalized standard deviation of the mass on the horizontal axis. The normalized standard deviation of the product of the perimeter and the sum of the covalent radius of atomic neighbors is shown on the vertical axis.

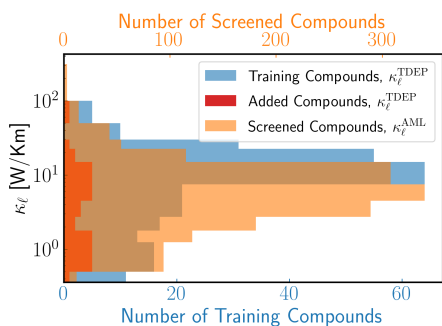


FIG. 9: Distribution of $\kappa_{\ell}^{\text{TDEP}}$ for training compounds and actively sampled compounds (bottom horizontal axis) and $\kappa_{\ell}^{\text{AML}}$ for screened compounds (top horizontal axis).

states (DOS) for Na_2TlSb at $T = 300$ K. There are two frequency gaps in the phonon dispersions (panel (a)), between acoustic branches and low-energy optical branches, and between low-energy optical branches and high-energy optical branches. There is a flat acoustic mode (blue in panel (a)) in the $U|K - \Gamma$ direction. This mode corresponds to low-velocity phonons with lifetimes on the higher end (blue markers in panel (b)). The combination of low phonon velocity and high lifetimes can yield low κ_{ℓ} seen from Eq. 3. Low-energy acoustic phonons (green and orange) have the highest phonon lifetimes. These phonons are reflected in the spectral $\kappa_{\ell}^{\text{TDEP}}$ (panel (c)) which peaks in this energy range. Acoustic phonons in the energy range 0.5-5.0 meV contribute $\sim 75\%$ to $\kappa_{\ell}^{\text{TDEP}}$. Interestingly, there is also a contribution to $\kappa_{\ell}^{\text{TDEP}}$ from low energy optical phonons, corresponding to an increase in the spectral $\kappa_{\ell}^{\text{TDEP}}$ at ~ 9 meV. A comparison of the phonon dispersions and DOS (panel (d)) reveals that the different atoms vibrate at different energies separated by the phonon gaps – a

TABLE I: Compounds with lowest κ_ℓ^{AML} at 300 K with $E_g^{\text{MP}} > 0$ eV from the MATERIALS PROJECT database [48].

Formula	mp-id	κ_ℓ^{AML} [W/Km]	E_g^{MP} [eV]
Ba ₂ AgSb	mp-984720	0.47	0.21
Ba ₂ AgBi	mp-1183207	0.50	0.06
Ba₂SnHg	mp-867912	0.56	0.04
Na₂TlSb	mp-866132	0.60	0.38
CsRb₂Bi	mp-1185546	0.64	0.19
CsK₂Bi	mp-581024	0.67	0.86
KBi₂	mp-23279	0.68	0.14
K ₂ RbBi	mp-1184754	0.70	0.14
CsRb₂Sb	mp-984761	0.73	0.65
K₂RbSb	mp-976148	0.80	0.69
Ba ₂ AsAu	mp-861937	0.81	0.52
K ₃ Bi	mp-568516	0.86	0.13
Sr ₂ BiAu	mp-867193	0.87	0.39
TlI	mp-23197	0.91	0.43
Rb ₃ Sb	mp-33018	0.96	1.48
Sr₂SnHg	mp-867169	0.98	1.81
RbGeI ₃	mp-571458	1.01	0.55
K ₃ Sb	mp-10159	1.01	0.68
CsK₂As	mp-1183938	1.02	0.71
Sr ₂ HgPb	mp-867207	1.05	0.12
CsRbAu₂	mp-1183931	1.06	0.71
CsCdBr ₃	mp-570231	1.09	0.70
Ca ₂ AsAu	mp-867113	1.11	0.20
RbAu	mp-30373	1.13	0.37
Sr₂HgGe	mp-1187086	1.13	0.05
CsCaBr ₃	mp-30056	1.13	4.39
RbCaBr₃	mp-1209227	1.14	4.08
KRbAu₂	mp-1184997	1.16	0.09
CsPbBr ₃	mp-600089	1.19	1.78
KI	mp-22901	1.22	3.24
Sr₂ZnPb	mp-1187098	1.25	0.05
CsSnBr ₃	mp-27214	1.26	0.60
NaTl₂Bi	mp-865145	1.26	0.20
Cs₂HgTe	mp-1185537	1.30	1.80

possible consequence of the large mass variation. The blue line in the DOS shows that Tl vibrations dominate in the acoustic range, while Sb contributes to low-energy optical phonons (orange), and Na to high-energy optical phonons (purple).

Ca₃AsBr₃ is the first compound selected during active sampling. It has seven atoms in the primitive cell, space group 221, $E_g = 1.67$ eV, and relatively low $\kappa_\ell^{\text{TDEP}} = 3.8$ W/Km. Fig. 10 panels (e)-(h) show the phonon dispersions, lifetimes, spectral $\kappa_\ell^{\text{TDEP}}$, and phonon DOS. The high lifetimes above 10^{-10} s correspond to the two low-energy acoustic branches (blue and orange). A substantial contribution to $\kappa_\ell^{\text{TDEP}}$ of $\sim 55\%$ originates from nine low to medium energy optical branches in the range 7.5-17.5 meV. The DOS shows that phonons in the acoustic region originate from Br vibrations, while high-energy optical phonons stem from Ca vibrations. Although Br and As have similar mass, As contributes less in the acoustic region. For this compound, unlike for Na₂TlSb, the different atoms vibrate in a wide range of energies,

possibly because of the relatively low mass variation in Ca₃AsBr₃.

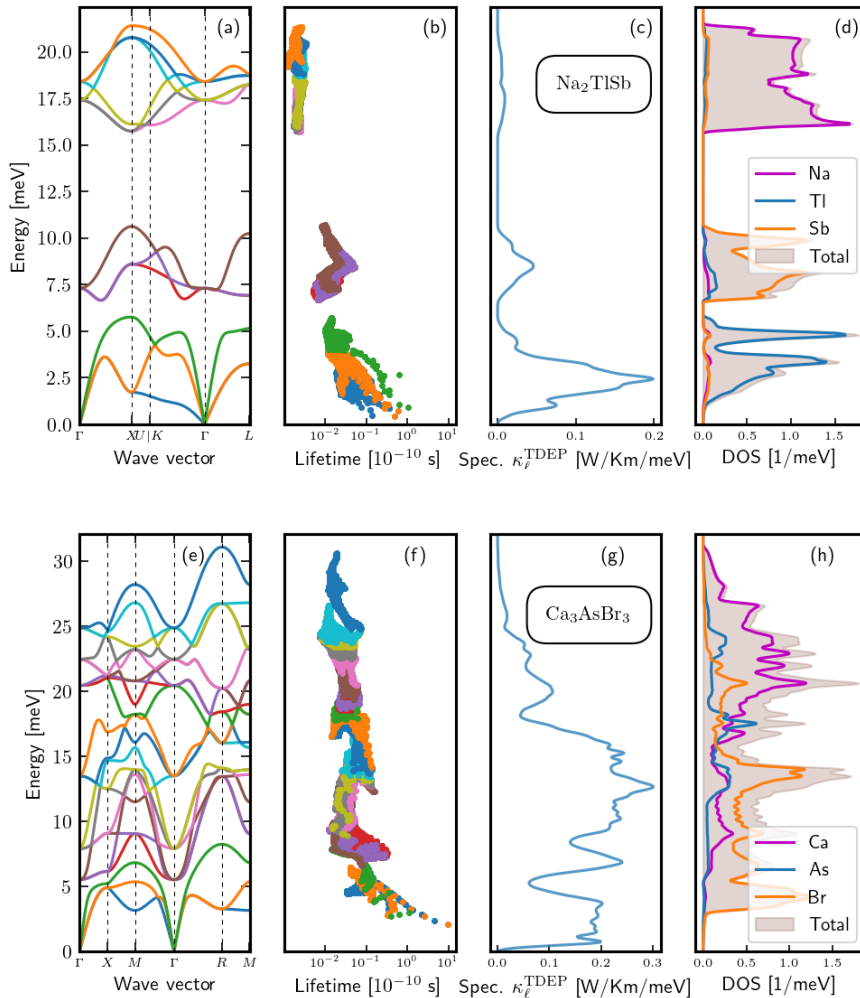


FIG. 10: (a) Phonon dispersions, (b) phonon lifetimes, (c) spectral $\kappa_{\ell}^{\text{TDEP}}$, and (d) site-projected DOS for Na_2TlSb . The coloring in panels (a) and (b) corresponds to phonon branches. The coloring in panel (d) corresponds to atoms. In (d) the PDOS for equivalent atoms is summed. (e)-(h) are the same plots for Ca_3AsBr_3 .

IV. SUMMARY AND CONCLUSION

In this paper a machine learning (ML) model was trained to predict the lattice thermal conductivity, κ_{ℓ} , of 1573 cubic compounds in the MATERIALS PROJECT database. The computed κ_{ℓ} used for training was obtained using density functional theory (DFT) and the temperature-dependent effective potential method.

Rather simple features based on atomic properties and the Voronoi structure were used. No DFT calculations were needed to obtain the features. We first demonstrated that active sampling based on uncertainties from Gaussian process regression performed better than random sampling. The active sampling model achieved accurate predictions with fewer training compounds; a crucial component in ML for the computationally expensive κ_{ℓ} . A total of 30 compounds were actively sampled from the

MATERIALS PROJECT and added to the training set. The model identified 34 low κ_ℓ compounds (≤ 1.3 W/Km) with band gaps that could be explored for thermoelectric applications. The full-Heusler Na_2TlSb achieved low ML predicted $\kappa_\ell = 0.6$ W/Km. To validate the prediction we calculated the κ_ℓ using DFT and the temperature-dependent effective potential method and obtained a $\kappa_\ell = 0.5$ W/Km. The low phonon velocities and phonon lifetimes in Na_2TlSb were shown to contribute to the low κ_ℓ .

Active sample selection enables the expansion of training sets to include other compound classes straightforwardly. It is also possible to include compounds with sublattice substitutions to expand the compound space to include more promising low κ_ℓ candidates. Model building and exploration can be executed with on-the-fly generation of prediction uncertainties, serving as a guideline pointing to compounds to sample next. As the active sampling scheme attempts to identify compounds dissimilar to the ones in the training set, it could be likely that dynamically unstable compounds are included. An em-

pirical or ML model predicting dynamical stability would be a complement to the selection scheme, enabling robust compound exploration in future studies yielding low κ_ℓ compounds dissimilar to typical ones.

ACKNOWLEDGMENTS

The computations in this work were funded by The Norwegian e-infrastructure for research and education, Sigma2, through grants No. nn9711k and nn2615k. This work is part of the Allotherm project (Project No. 314778) supported by the Research Council of Norway.

Appendix A: Machine Learning Features

Appendix B: Calculated Lattice Thermal Conductivity

-
- [1] G. Snyder and E. Toberer, Complex thermoelectric materials, *Nat. Mater.* **7**, 105 (2008).
 - [2] Z. Soleimani, S. Zoras, B. Ceranic, S. Shahzad, and Y. Cui, A review on recent developments of thermoelectric materials for room-temperature applications, *Sustain. Energy Technol. Assess.* **37**, 100604 (2020).
 - [3] O. Farhat, J. Faraj, F. Hachem, C. Castelain, and M. Khaled, A recent review on waste heat recovery methodologies and applications: Comprehensive review, critical analysis and potential recommendations, *Cleaner Eng. Technol.* **6**, 100387 (2022).
 - [4] K. Berland, N. Shulumba, O. Hellman, C. Persson, and O. M. Løvvik, Thermoelectric transport trends in group 4 half-Heusler alloys, *J. Appl. Phys.* **126**, 145102 (2019).
 - [5] M. W. Gaultois and T. D. Sparks, How much improvement in thermoelectric performance can come from reducing thermal conductivity?, *Appl. Phys. Lett.* **104**, 113906 (2014).
 - [6] D. M. Rowe and V. S. Shukla, The effect of phonon-grain boundary scattering on the lattice thermal conductivity and thermoelectric conversion efficiency of heavily doped fine-grained, hot-pressed silicon germanium alloy, *J. Appl. Phys.* **52**, 7421 (1981).
 - [7] M. Schrade, K. Berland, S. Eliassen, M. Guzik, C. Echevarria-Bonet, M. Sørby, P. Jenus, B. Hauback, R. Tofan, A. Gunnæs, C. Persson, O. Løvvik, and T. Finstad, The role of grain boundary scattering in reducing the thermal conductivity of polycrystalline XNiSn ($\text{X} = \text{Hf, Zr, Ti}$) half-Heusler alloys, *Sci. Rep.* **7** (2017).
 - [8] Y. Zheng, T. J. Slade, L. Hu, X. Y. Tan, Y. Luo, Z.-Z. Luo, J. Xu, Q. Yan, and M. G. Kanatzidis, Defect engineering in thermoelectric materials: what have we learned?, *Chem. Soc. Rev.* **50**, 9022 (2021).
 - [9] C. Zhao, Z. Li, T. Fan, C. Xiao, and Y. Xie, Defects engineering with multiple dimensions in thermoelectric materials, *Research* **2020**, 1 (2020).
 - [10] J. Zhou, L. Huang, Z. Wang, Z. Liu, W. Cai, and J. Sui, Effect of Cd isoelectronic substitution on thermoelectric properties of $\text{Zn}_{0.995}\text{Na}_{0.005}\text{Sb}$, *J. Materiomics* **2**, 324 (2016).
 - [11] J. Zhang, L. Huang, C. Zhu, C. Zhou, B. Jabar, J. Li, X. Zhu, L. Wang, C. Song, S. Xin, D. Li, and X.-y. Qin, Design of domain structure and realization of ultralow thermal conductivity for record-high thermoelectric performance in chalcopyrite, *Adv. Mater.* **31** (2019).
 - [12] S. N. H. Eliassen, A. Katre, G. K. H. Madsen, C. Persson, O. M. Løvvik, and K. Berland, Lattice thermal conductivity of $\text{Ti}_x\text{Zr}_y\text{Hf}_{1-x-y}\text{NiSn}$ half-Heusler alloys calculated from first principles: Key role of nature of phonon modes, *Phys. Rev. B* **95**, 045202 (2017).
 - [13] R. Tranàs, O. M. Løvvik, and K. Berland, Attaining low lattice thermal conductivity in half-Heusler sublattice solid solutions: Which substitution site is most effective?, *Electron. Mater.* **3**, 1 (2022).
 - [14] J. Yu, C. Fu, Y. Liu, K. Xia, U. Aydemir, T. Chasapis, G. Snyder, X. Zhao, and T. Zhu, Unique role of refractory Ta alloying in enhancing the figure of merit of NbFeSb thermoelectric materials, *Adv. Energy Mater.* **8**, 1701313 (2017).
 - [15] T. Luo, F. Serrano-Sánchez, H. Bishara, S. Zhang, R. Bueno Villoro, J. J. Kuo, C. Felser, C. Scheu, G. J. Snyder, J. P. Best, G. Dehm, Y. Yu, D. Raabe, C. Fu, and B. Gault, Dopant-segregation to grain boundaries controls electrical conductivity of n-type NbCo(Pt)Sn half-Heusler alloy mediating thermoelectric performance, *Acta Mater.* **217**, 117147 (2021).
 - [16] L. Hu, T. Zhu, X. Liu, and X. Zhao, Point defect engineering of high-performance bismuth-telluride-based thermoelectric materials, *Adv. Funct. Mater.* **24**, 5211 (2014).
 - [17] C.-H. Su, Experimental determination of lattice thermal conductivity and Lorenz number as functions of temperature for n-type PbTe , *Mater. Today Phys.* **5**, 58 (2018).

TABLE II: Features used for training the GPR model. The features obtained from feature selection are marked by x in the two rightmost columns. The feature importances are shown in parenthesis. Further information is detailed in Sec. IIC

Feature	Description	SFFS before ac- tive sampling	SFFS after ac- tive sampling
$\langle r \rangle$	Average covalent radius	x (0.098)	x (0.232)
$\bar{m}(r)$	Median covalent radius	x (0.066)	x (0.182)
$\sigma(r)$	Standard deviation covalent radius	-	x (0.045)
$\langle \chi \rangle$	Average electronegativity	-	-
$\bar{m}(\chi)$	Median electronegativity	-	-
$\sigma(\chi)$	Standard deviation electronegativity	x (0.328)	x (0.126)
$\langle \mu_{\text{dip}} \rangle$	Average dipole polarizability	x (0.097)	x (0.060)
$\bar{m}(\mu_{\text{dip}})$	Median dipole polarizability	x (0.045)	-
$\sigma(\mu_{\text{dip}})$	Standard deviation dipole polarizability	-	-
$\langle m \rangle$	Average atomic mass	-	-
$\bar{m}(m)$	Median atomic mass	x (0.034)	-
$\sigma(m)$	Standard deviation atomic mass	-	-
$\langle N_{\text{val}} \rangle$	Average number of valence electrons	x (0.092)	x (0.062)
$\bar{m}(N_{\text{val}})$	Median number of valence electrons	x (0.187)	x (0.146)
$\sigma(N_{\text{val}})$	Standard deviation number of valence electrons	-	-
$\langle N_{\text{s}} \rangle$	Average number of s electrons	-	-
$\bar{m}(N_{\text{s}})$	Median number of s electrons	-	x (0.037)
$\sigma(N_{\text{s}})$	Standard deviation number of s electrons	-	-
$\langle N_{\text{p}} \rangle$	Average number of p electrons	-	-
$\bar{m}(N_{\text{p}})$	Median number of p electrons	x (0.066)	-
$\sigma(N_{\text{p}})$	Standard deviation number of p electrons	-	-
$\langle N_{\text{d}} \rangle$	Average number of d electrons	-	-
$\bar{m}(N_{\text{d}})$	Median number of d electrons	-	-
$\sigma(N_{\text{d}})$	Standard deviation number of d electrons	-	-
$\langle N_{\text{f}} \rangle$	Average number of f electrons	-	-
$\bar{m}(N_{\text{f}})$	Median number of f electrons	-	-
$\sigma(N_{\text{f}})$	Standard deviation number of f electrons	-	-
N_{ele}	Number of elements	x (0.048)	-
$\langle \sum_n L_{\text{per},n} \rangle$	Average sum of face perimeters of Voronoi polyhedrons	-	x (0.044)
$\bar{m}(\sum_n L_{\text{per},n})$	Median sum of face perimeters of Voronoi polyhedrons	x (0.051)	x (0.053)
$\sigma(\sum_n L_{\text{per},n})$	Standard deviation sum of face perimeters of Voronoi polyhedrons	x (0.058)	-
$\langle V_{\text{covrad}}/V_{\text{Vor}} \rangle$	Average volume of sphere defined by covalent radius divided by the volume of the Voronoi polyhedron	x (0.069)	x (0.127)
$\bar{m}(V_{\text{covrad}}/V_{\text{Vor}})$	Median volume of sphere defined by covalent radius divided by the volume of the Voronoi polyhedron	x (0.033)	x (0.027)
$\sigma(V_{\text{covrad}}/V_{\text{Vor}})$	Standard deviation volume of sphere defined by covalent radius divided by the volume of the Voronoi polyhedron	-	x (0.045)
$\langle \sum_n L_{\text{per},n}(r+r_n) \rangle$	Average product of face perimeter and sum of covalent radius of atom neighbors	x (0.032)	x (0.082)
$\bar{m}(\sum_n L_{\text{per},n}(r+r_n))$	Median product of face perimeter and sum of covalent radius of atom neighbors	-	-
$\sigma(\sum_n L_{\text{per},n}(r+r_n))$	Standard deviation product of face perimeter and sum of covalent radius of atom neighbors	x (0.146)	x (0.092)
$\langle V_{\text{Vor}} \rangle$	Average Voronoi polyhedron volume	-	x (0.004)
$\bar{m}(V_{\text{Vor}})$	Median Voronoi polyhedron volume	-	-
$\sigma(V_{\text{Vor}})$	Standard deviation Voronoi polyhedron volume	-	-
$\langle (\sum_n L_{\text{per},n})^2 / \sum_n L_{\text{per},n}^2 \rangle$	Average effective coordination number	x (0.071)	x (0.093)
$\bar{m}((\sum_n L_{\text{per},n})^2 / \sum_n L_{\text{per},n}^2)$	Median effective coordination number	x (0.060)	-
$\sigma((\sum_n L_{\text{per},n})^2 / \sum_n L_{\text{per},n}^2)$	Standard deviation effective coordination number	x (0.054)	x (0.140)
$\langle N_{\text{cor}} \rangle$	Average coordination number	-	-
$\bar{m}(N_{\text{cor}})$	Median coordination number	-	-
$\sigma(N_{\text{cor}})$	Standard deviation coordination number	-	-
S_1	1st Steinhardt's parameter	-	-
S_2	2nd Steinhardt's parameter	x (0.030)	-
S_3	3rd Steinhardt's parameter	-	-
S_4	4th Steinhardt's parameter	-	-

TABLE III: Calculated $\kappa_\ell^{\text{TDEP}}$ at 300 K for 268 compounds used in training. The actively sampled compounds are shown in bold.

κ_ℓ [W/Km]	κ_ℓ [W/Km]	κ_ℓ [W/Km]	κ_ℓ [W/Km]	κ_ℓ [W/Km]	κ_ℓ [W/Km]				
Ag ₂ S	0.19	CsBr	2.66	HfPdPb	9.33	TaCoPb	15.60	TiPdGe	23.11
CsRb ₃	0.20	CuCl	2.69	TaOsBi	9.33	HfPtSn	15.70	VRuAs	23.19
KRb ₃	0.24	Rb ₂ O	2.80	SrF ₂	9.54	HfPdGe	15.70	TiIrAs	23.30
K ₃ Na	0.29	RbI	2.82	TaRhPb	9.61	ZrGeRu ₂	15.94	TaFeAs	24.09
CuCl	0.38	CsI	2.93	NbIrPb	9.74	HfNiPb	16.03	NbFeSb	24.10
Na ₂ TlBi	0.39	CaMg ₃	3.10	Li ₂ Te	9.90	TiSnRu ₂	16.08	TaOsSb	24.70
I₄Sn	0.41	RbF	3.16	ZrPdPb	10.05	ZrNiSn	16.18	NbOsAs	25.21
Ba ₂ BiAu	0.41	BaBiK	3.17	LaOF	10.14	TeRuZr	16.35	VCoGe	25.28
CsAu	0.45	Li ₂ TlBi	3.18	MgTe	10.24	ZrPdGe	16.40	TaCoGe	25.52
Ba ₂ HgPb	0.49	RbBr	3.19	KF	10.35	NbCoPb	16.41	TaRhGe	25.68
Cu ₂ S	0.50	AgI	3.20	ZrPtPb	10.37	HfGeRu ₂	16.47	GaNiNb	25.76
CsK ₂ Bi	0.53	LiPt₂	3.23	HfIrBi	10.44	F₃Mg₃N	16.50	NbIrGe	26.08
Ba ₂ SbAu	0.56	Li ₂ CaSi	3.41	CrGa₄	10.45	TiNiSn	16.56	TaRuAs	26.08
TlBr	0.56	Rb ₂ S	3.46	TiIrBi	10.58	VCoSn	17.32	AlAuHf	26.28
InTl₃	0.59	Cu ₃ As	3.48	NiZn	10.69	TaIrSn	17.51	GeFeW	26.50
AgI	0.61	Li ₂ InBi	3.59	Be₂Cu	11.02	TaRhSn	17.63	VOsAs	26.81
Bi₂Cs	0.79	Sr ₃ BiN	3.67	TaRuBi	11.08	NbOsSb	17.64	TaFeSb	27.05
Sr ₂ SbAu	0.82	LaBiPd	3.74	HfRhBi	11.10	ZrIrSb	17.65	VRhGe	27.12
MgHg ₃	0.83	Li₃S₄V	3.81	ZrIrBi	11.14	ZrRhSb	17.76	TaOsAs	27.20
TiCl	0.84	Ca₃AsBr₃	3.82	VIrSn	11.19	HfNiGe	17.77	VhGe	27.54
AgCl	0.93	YNiP	3.84	BiNiSc	11.37	HfRhSb	17.82	GaPtTa	27.93
Ca ₂ SbAu	0.94	AgCl	3.88	ZrIrAs	11.45	HfNiSn	17.85	TiRhAs	28.42
AgBr	0.96	LiBr	4.21	ScSbPt	11.51	NbIrSn	18.08	SiCoTa	28.79
Hg₄Ni	1.12	CdPNa	4.51	NbOsBi	11.64	NbFeBi	18.30	TiRGe	29.03
AgBr	1.13	RbF	4.60	TaTiO ₃	11.66	ZrNiGe	18.30	NbCoGe	29.24
MgZn₂	1.18	HgSe	4.64	ZrAsIr	11.70	HfPtGe	18.39	ReW₃	29.32
Al₂Ba	1.22	Ca ₂ CdPb	4.83	ZrRhBi	11.73	ZrCoBi	18.41	TiCoAs	29.49
KNa ₂ Bi	1.26	BaO	4.88	ScNiP	11.76	ZrSbRh	18.62	NbRuAs	29.65
Cs₂Se	1.35	LiCaAs	4.99	NbRhPb	11.83	TiTeRu	18.64	NbRhGe	29.85
YbPd	1.36	Au₂S	5.15	VRhSn	11.86	HfCoAs	18.84	TiCoBi	30.50
LaPtSb	1.37	CdF ₂	5.17	ScTeRh	12.41	ZnSe	18.87	VFeAs	30.58
RbCl	1.38	CePt ₃	5.20	ScSbPd	12.47	HfRhAs	18.88	VGaFe ₂	33.72
Ba₃Sr	1.50	Cu₃TaTe₄	5.23	Li ₂ Se	12.97	ZrSnRu ₂	18.92	LaAs	34.49
KNa ₂ Sb	1.53	Cu₂O	5.49	AlGeLi	13.00	TaFeBi	19.06	NbCoSi	34.80
Hg₄Pt	1.54	LaY ₃	5.55	LiCdAs	13.04	TiRhSb	19.07	NbFeAs	35.39
Cd ₃ Pd	1.71	RbBr	5.58	TiRhBi	13.12	NbRuSb	19.19	GaAs	38.31
LaRhTe	1.72	Ca ₂ Ge	5.72	CdS₄Zn₃	13.14	NbCoSn	19.51	VSiRh	42.88
CaLi₂	1.77	BrK	5.98	ZrPtGe	13.50	LaP	19.81	LiBSi	43.98
TiPdF ₃	1.77	LiZnSb	6.01	VRuSb	13.72	TiIrSb	19.89	AlVFe ₂	45.89
AuBe	1.78	Cu₃LiO₃	6.50	TiPdSn	13.80	HfCoBi	19.90	GeSi₇	59.48
TlZnF ₃	1.86	Ca ₃ SbN	7.02	NbRuBi	13.89	HfCoSb	19.93	AlBi	61.10
Pt ₃ Pb	2.02	TePb	7.07	TiPtSn	14.12	HfSnRu ₂	20.32	RuO ₂	78.94
CaCl ₂	2.04	HfBRh ₃	7.12	ZrPdSn	14.25	VFeSb	20.44	BeTe	89.09
Sr ₃ Ca	2.04	SnTe	7.20	VOsSb	14.34	TaRuSb	21.24	LiBeSb	92.45
NaNZnAs	2.06	GeTe	7.57	HfIrAs	14.43	HfIrSb	21.32	SiC	409.91
Ca ₂ HgPb	2.06	AuZr₃	8.52	ZrNiPb	14.45	TaCoSn	21.57	BP	442.90
BaPd₂	2.09	BiPdSc	8.61	AcAlO ₃	14.54	B₆Ca	22.01	BN	658.74
CsF	2.20	TiPtPb	8.71	HfPdSn	14.55	ZrCoSb	22.20	C	2039.96
RbI	2.22	TaIrPb	8.72	HfSiRu ₂	14.67	TiNiGe	22.22		
Rb ₂ Te	2.26	HfPtPb	8.75	VSbRu	14.76	TiNiPb	22.22		
BaZn	2.29	LiMgSb	9.09	ZrPtSn	14.80	ZrRhAs	22.23		
MoZn₆	2.44	BiNiY	9.12	AsNiSc	14.81	ZrCoAs	22.32		
CsI	2.45	TiPdPb	9.22	ZrSiRu ₂	14.91	TeFeTi	22.48		
Al ₃ Ga	2.59	Sc ₃ BPb	9.28	AlSiLi	15.28	TiCoSb	22.66		
InAg ₃	2.63	CuBr	9.31	NbRhSn	15.33	TiPtGe	22.74		

- [18] Y. Lu, T. Sun, and D.-B. Zhang, Lattice anharmonicity, phonon dispersion, and thermal conductivity of PbTe studied by the phonon quasiparticle approach, *Phys. Rev. B* **97** (2018).
- [19] C. B. Satterthwaite and R. W. Ure, Electrical and thermal properties of Bi_2Te_3 , *Phys. Rev.* **108**, 1164 (1957).
- [20] I. T. Witting, T. C. Chasapis, F. Ricci, M. Peters, N. A. Heinz, G. Hautier, and G. J. Snyder, The thermoelectric properties of bismuth telluride, *Adv. Electron. Mater.* **5**, 1800904 (2019).
- [21] C. Zhou, Y. K. Lee, Y. Yu, S. Byun, Z.-Z. Luo, H. Lee, B. Ge, Y.-L. Lee, X. Chen, J. Y. Lee, O. Cojocaru-Mirédin, H. Chang, J. Im, S.-P. Cho, M. Wuttig, V. P. Dravid, M. G. Kanatzidis, and I. Chung, Polycrystalline SnSe with a thermoelectric figure of merit greater than the single crystal, *Nat. Mater.* **20**, 10.1038/s41563-021-01064-6 (2021).
- [22] Y. Xiao, C. Chang, Y. Pei, D. Wu, K. Peng, X. Zhou, S. Gong, J. He, Y. Zhang, Z. Zeng, and L.-D. Zhao, Origin of low thermal conductivity in SnSe, *Phys. Rev. B* **94**, 125203 (2016).
- [23] O. Caballero-Calero, J. R. Ares, and M. Martín-González, Environmentally friendly thermoelectric materials: High performance from inorganic components with low toxicity and abundance in the earth, *Adv. Sustain. Syst.* **5**, 2100095 (2021).
- [24] P. Ren, Y. Liu, J. He, T. Lv, J. Gao, and G. Xu, Recent advances in inorganic material thermoelectrics, *Inorg. Chem. Front.* **5**, 2380 (2018).
- [25] L. Chen, H. Tran, R. Batra, C. Kim, and R. Ramprasad, Machine learning models for the lattice thermal conductivity prediction of inorganic materials, *Comput. Mater. Sci.* **170**, 109155 (2019).
- [26] L. M. Antunes, Vikram, J. J. Plata, A. V. Powell, K. T. Butler, and R. Grau-Crespo, Machine learning approaches for accelerating the discovery of thermoelectric materials, in *Machine Learning in Materials Informatics: Methods and Applications*, Chap. 1, pp. 1–32.
- [27] P. Gorai, V. Stevanovic, and E. Toberer, Computationally guided discovery of thermoelectric materials, *Nat. Rev. Mater.* **2**, 17503 (2017).
- [28] Y. Gan, G. Wang, J. Zhou, and Z. Sun, Prediction of thermoelectric performance for layered IV-VI semiconductors by high-throughput ab initio calculations and machine learning, *npj Comput. Mater.* **7**, 10.1038/s41524-021-00645-y (2021).
- [29] T. Jia, Z. Feng, S. Guo, X. Zhang, and Y. Zhang, Screening promising thermoelectric materials in binary chalcogenides through high-throughput computations, *ACS Appl. Mater. Interfaces.* **12**, 11852 (2020), pMID: 32069390.
- [30] J. Li, W. Hu, and J. Yang, High-throughput screening of rattling-induced ultralow lattice thermal conductivity in semiconductors, *J. Am. Chem. Soc.* **144**, 4448 (2022), pMID: 35230828.
- [31] K. Pal, Y. Xia, J. Shen, J. He, Y. Luo, M. G. Kanatzidis, and C. Wolverton, Accelerated discovery of a large family of quaternary chalcogenides with very low lattice thermal conductivity, *npj Comput. Mater.* **7**, 82 (2021).
- [32] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling, *Phys. Rev. X* **4** (2014).
- [33] Z. Feng, Y. Fu, Y. Zhang, and D. J. Singh, Characterization of rattling in relation to thermal conductivity: ordered half-Heusler semiconductors, *Phys. Rev. B* **101**, 064301 (2020).
- [34] A. Togo, L. Chaput, and I. Tanaka, Distributions of phonon lifetimes in Brillouin zones, *Phys. Rev. B* **91**, 094306 (2015).
- [35] O. Hellman and I. A. Abrikosov, Temperature-dependent effective third-order interatomic force constants from first principles, *Phys. Rev. B* **88**, 144301 (2013).
- [36] F. Zhou, W. Nielson, Y. Xia, and V. Ozoliņš, Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations, *Phys. Rev. Lett.* **113**, 185501 (2014).
- [37] F. Eriksson, E. Fransson, and P. Erhart, The hiphive package for the extraction of high-order force constants by machine learning, *Adv. Theory Simul.* **2**, 1800184 (2019).
- [38] M. Li, L. Dai, and Y. Hu, Machine learning for harnessing thermal energy: From materials discovery to system optimization, *ACS Energy Lett.* **0**, 3204 (0).
- [39] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization, *Phys. Rev. Lett.* **115**, 205901 (2015).
- [40] X. Wang, Y. Zhao, S. Zeng, Z. Wang, Y. Chen, and J. Ni, Cubic halide perovskites as potential low thermal conductivity materials: A combined approach of machine learning and first-principles calculations, *Phys. Rev. B* **105**, 014310 (2022).
- [41] R. Juneja, G. Yumnam, S. Satsangi, and A. K. Singh, Coupling the high-throughput property map to machine learning for predicting lattice thermal conductivity, *Chem. Mater.* **31**, 5145 (2019).
- [42] X. Wang, S. Zeng, Z. Wang, and J. Ni, Identification of crystalline materials with ultra-low thermal conductivity based on machine learning study, *The Journal of Physical Chemistry C* **124**, 8488 (2020).
- [43] R. Jaafreh, Y. S. Kang, and K. Hamad, Lattice thermal conductivity: An accelerated discovery guided by machine learning, *ACS Appl. Mater. Interfaces* **13**, 57204 (2021).
- [44] K. Pal, C. Park, Y. Xia, J. Shen, and C. Wolverton, Scale-invariant machine-learning model accelerates the discovery of quaternary chalcogenides with ultralow lattice thermal conductivity, *npj Comput. Mater.* **8**, 10.1038/s41524-022-00732-8 (2022).
- [45] J. M. Choi, K. Lee, S. Kim, M. Moon, W. Jeong, and S. Han, Accelerated computation of lattice thermal conductivity using neural network interatomic potentials, *Comput. Mater. Sci.* **211**, 111472 (2022).
- [46] O. M. Løvvik, E. Flage-Larsen, and G. Skomedal, Screening of thermoelectric silicides with atomistic transport calculations, *J. Appl. Phys.* **128**, 125105 (2020).
- [47] R. Juneja and A. K. Singh, Guided patchwork kriging to develop highly transferable thermal conductivity prediction models, *J. Phys. Mater.* **3**, 024006 (2020).
- [48] 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, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.* **1**, 011002

- (2013).
- [49] J. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD), *JOM* **65** (2013).
- [50] C. Draxl and M. Scheffler, The NOMAD laboratory: from data sharing to artificial intelligence, *JPhys Materials* **2**, 036001 (2019).
- [51] Y. Zhang and C. Ling, A strategy to apply machine learning to small datasets in materials science, *npj Comput. Mater.* **4**, 10.1038/s41524-018-0081-z (2018).
- [52] R. Tranås, O. M. Løvvik, O. Tomic, and K. Berland, Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis, *Comput. Mater. Sci.* **202**, 110938 (2022).
- [53] H. Wei, H. Bao, and X. Ruan, Perspective: Predicting and optimizing thermal transport properties with machine learning methods, *Energy and AI* **8**, 100153 (2022).
- [54] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, On-the-fly active learning of interpretable bayesian force fields for atomistic rare events, *npj Comput. Mater.* **6**, 1 (2020).
- [55] M. K. Bisbo and B. Hammer, Efficient global structure optimization with a machine-learned surrogate model, *Phys. Rev. Lett.* **124**, 086102 (2020).
- [56] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with bayesian inference, *Phys. Rev. Lett.* **122**, 225701 (2019).
- [57] D. Schwalbe-Koda, A. Tan, and R. Gómez-Bombarelli, Differentiable sampling of molecular geometries with uncertainty-based adversarial attacks, *Nat. Commun.* **12** (2021).
- [58] O. Hellman, I. A. Abrikosov, and S. I. Simak, Lattice dynamics of anharmonic solids from first principles, *Phys. Rev. B* **84**, 180301 (2011).
- [59] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Gaussian process regression for materials and molecules, *Chem. Rev.* **121**, 10073 (2021).
- [60] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, and I. Tanaka, Representation of compounds for machine-learning prediction of physical properties, *Phys. Rev. B* **95**, 144110 (2017).
- [61] M. J. Burn and P. L. A. Popelier, Creating gaussian process regression models for molecular simulations using adaptive sampling, *J. Chem. Phys.* **153**, 054111 (2020).
- [62] M. M. Noack, G. S. Doerk, R. Li, J. K. Streit, R. A. Vaia, K. G. Yager, and M. Fukuto, Autonomous materials discovery driven by gaussian process regression with inhomogeneous measurement noise and anisotropic kernels, *Sci. Rep.* **10**, 10.1038/s41598-020-74394-1 (2020).
- [63] B. Goodarzi and A. R. Bahramian, Applying machine learning for predicting thermal conductivity coefficient of polymeric aerogels, *J. Therm. Anal. Calorim.* , 1 (2021).
- [64] X. Yue, Y. Wen, J. H. Hunt, and J. Shi, Active learning for gaussian process considering uncertainties with application to shape control of composite fuselage, *IEEE Trans. Autom. Sci. Eng.* **18**, 36 (2021).
- [65] B. Cordero, V. Gómez, A. Platero-Prats, M. Revés, J. Echeverria, E. Cremades, F. Barragán, and S. Alvarez, Covalent radii revisited, *Dalton trans.* **21**, 2832 (2008).
- [66] L. C. Allen, Electronegativity is the average one-electron energy of the valence-shell electrons in ground-state free atoms, *J. Am. Chem. Soc.* **111**, 9003 (1989).
- [67] J. Meija, T. Coplen, M. Berglund, W. Brand, P. De Bièvre, M. Gröning, N. Holden, J. Irrgeher, R. Loss, T. Walczyk, and T. Prohaska, Atomic weights of the elements 2013 (iupac technical report), *Pure Appl. Chem.* **88**, 265 (2016).
- [68] mendeleev – a python resource for properties of chemical elements, ions and isotopes, ver. 0.10.0, <https://github.com/lmmentel/mendeleev> (2014-).
- [69] P. Schwerdtfeger and J. K. Nagle, 2018 table of static dipole polarizabilities of the neutral elements in the periodic table, *Mol. Phys.* **117**, 1200 (2019).
- [70] G. Voronoi, Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième mémoire. recherches sur les paralléloèdres primitifs., *Journal für die reine und angewandte Mathematik (Crelles Journal)* **1908**, 198 (1908).
- [71] L. Ward, R. Liu, A. Krishna, V. I. Hegde, A. Agrawal, A. Choudhary, and C. Wolverton, Including crystal structure attributes in machine learning models of formation energies via voronoi tessellations, *Phys. Rev. B* **96**, 024104 (2017).
- [72] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Bond-orientational order in liquids and glasses, *Phys. Rev. B* **28**, 784 (1983).
- [73] W. Mickel, S. Kapfer, G. Schröder-Turk, and K. Mecke, Shortcomings of the bond orientational order parameters for the analysis of disordered particulate matter, *J. Chem. Phys.* **138**, 044501 (2013).
- [74] S. Menon, G. D. Leines, and J. Rogal, pycs: A python module for structural analysis of atomic environments, *Journal of Open Source Software* **4**, 1824 (2019).
- [75] F. Knoop, T. A. R. Purcell, M. Scheffler, and C. Carbogno, Anharmonicity measure for materials, *Phys. Rev. Mater.* **4**, 083809 (2020).
- [76] F. Legrain, J. Carrete, A. van Roekeghem, G. K. Madsen, and N. Mingo, Materials screening for the discovery of new half-Heuslers: Machine learning versus ab initio methods, *J. Phys. Chem. B* **122**, 625 (2018), pMID: 28742351.
- [77] R. Jaafreh, K. Yoo Seong, J.-G. Kim, and K. Hamad, A deep learning perspective into the figure-of-merit of thermoelectric materials, *Mater. Lett.* **319**, 132299 (2022).
- [78] P. Pudil, J. Novovičová, and J. Kittler, Floating search methods in feature selection, *Pattern Recognit. Lett.* **15**, 1119 (1994).
- [79] S. Raschka, MLxtend: Providing machine learning and data science utilities and extensions to Python’s scientific computing stack, *JOSS* **3**, 638 (2018).
- [80] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, and D. Cournapeau, Scikit-learn: machine learning in Python, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [81] B. Wei, X. Yu, C. Yang, X. Rao, X. Wang, S. Chi, X. Sun, and J. Hong, Low-temperature anharmonicity and the thermal conductivity of Cesium Iodide, *Phys. Rev. B* **99**, 184301 (2019).

- [82] S. Mukhopadhyay, D. Bansal, O. Delaire, D. Perrodin, E. Bourret-Courchesne, D. J. Singh, and L. Lindsay, The curious case of cuprous chloride: Giant thermal resistance and anharmonic quasiparticle spectra driven by dispersion nesting, *Phys. Rev. B* **96**, 100301 (2017).
- [83] S. Ju, R. Yoshida, C. Liu, S. Wu, K. Hongo, T. Tadano, and J. Shiomi, Exploring diamond-like lattice thermal conductivity crystals via feature-based transfer learning, *Phys. Rev. Mater.* **5**, 053801 (2021).
- [84] T. Pandey, L. Lindsay, B. C. Sales, and D. S. Parker, Lattice instabilities and phonon thermal transport in TlBr, *Phys. Rev. Mater.* **4**, 045403 (2020).
- [85] Y. Yao, B.-P. Zhang, J. Pei, Y.-C. Liu, and J.-F. Li, Thermoelectric performance enhancement of Cu_2S by Se doping leading to a simultaneous power factor increase and thermal conductivity reduction, *J. Mater. Chem. C* **5**, 7845 (2017).
- [86] H. Chen, Z. Yue, D. Ren, H. Zeng, T. Wei, K. Zhao, R. Yang, P. Qiu, L. Chen, and X. Shi, Thermal conductivity during phase transitions, *Adv. Mater.* , 1806518 (2018).
- [87] J. He, M. Amsler, Y. Xia, S. S. Naghavi, V. I. Hegde, S. Hao, S. Goedecker, V. Ozoliņš, and C. Wolverton, Ultralow thermal conductivity in full-Heusler semiconductors, *Phys. Rev. Lett.* **117**, 046602 (2016).
- [88] J. Park, Y. Xia, and V. Ozoliņš, High thermoelectric power factor and efficiency from a highly dispersive band in Ba_2BiAu , *Phys. Rev. Appl.* **11**, 014058 (2019).
- [89] W. Wang, Z. Dai, X. Wang, Q. Zhong, Y. Zhao, and S. Meng, Low lattice thermal conductivity and high figure of merit in n-type doped full-Heusler compounds X_2YAu ($\text{X}=\text{Sr}, \text{Ba}$; $\text{Y}=\text{As}, \text{Sb}$), *Int. J. Energy Res.* **45**, 20949 (2021).
- [90] S. P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, and K. A. Persson, The materials application programming interface (api): A simple, flexible and efficient api for materials data based on representational state transfer (rest) principles, *Comput. Mater. Sci.* **97**, 209 (2015).
- [91] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, *Phys. Rev. B* **47**, 558 (1993).
- [92] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [93] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* **6**, 15 (1996).
- [94] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [95] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán, Assessing the performance of recent density functionals for bulk solids, *Phys. Rev. B* **79**, 155107 (2009).
- [96] K. Berland and P. Hyldgaard, Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional, *Phys. Rev. B* **89**, 035412 (2014).
- [97] K. Berland, C. A. Arter, V. R. Cooper, K. Lee, B. I. Lundqvist, E. Schröder, T. Thonhauser, and P. Hyldgaard, van der Waals density functionals built upon the electron-gas tradition: Facing the challenge of competing interactions, *J. Chem. Phys.* **140**, 18A539 (2014).
- [98] L. He, F. Liu, G. Hautier, M. J. T. Oliveira, M. A. L. Marques, F. D. Vila, J. J. Rehr, G.-M. Rignanesse, and A. Zhou, Accuracy of generalized gradient approximation functionals for density-functional perturbation theory calculations, *Phys. Rev. B* **89**, 064305 (2014).
- [99] C. M. Frostenson, E. J. Granhed, V. Shukla, P. A. T. Olsson, E. Schröder, and P. Hyldgaard, Hard and soft materials: putting consistent van der waals density functionals to work, *Electron. Struct.* **4**, 014001 (2022).
- [100] D. O. Lindroth and P. Erhart, Thermal transport in van der waals solids from first-principles calculations, *Phys. Rev. B* **94**, 115205 (2016).
- [101] L. Gharaee, P. Erhart, and P. Hyldgaard, Finite-temperature properties of nonmagnetic transition metals: Comparison of the performance of constraint-based semilocal and nonlocal functionals, *Phys. Rev. B* **95**, 085147 (2017).
- [102] N. Shulumba, O. Hellman, and A. J. Minnich, Intrinsic localized mode and low thermal conductivity of PbSe , *Phys. Rev. B* **95**, 014302 (2017).
- [103] O. Anderson, A simplified method for calculating the debye temperature from elastic constants, *J. Phys. Chem. Solids* **24**, 909 (1963).
- [104] M. K. Jana and K. Biswas, Crystalline solids with intrinsically low lattice thermal conductivity for thermoelectric energy conversion, *ACS Energy Lett.* **3**, 1315 (2018).
- [105] J. S. Kang, M. Li, H. Wu, H. Nguyen, and Y. Hu, Experimental observation of high thermal conductivity in boron arsenide, *Science* **361**, 575 (2018).
- [106] J. S. Kang, M. Li, H. Wu, H. Nguyen, and Y. Hu, Experimental observation of high thermal conductivity in boron arsenide, *Science* **361**, 575 (2018).
- [107] K. Berland, O. M. Løvvik, and R. Tranås, Discarded gems: Thermoelectric performance of materials with band gap emerging at the hybrid-functional level, *Appl. Phys. Lett.* **119**, 081902 (2021).
- [108] S.-F. Wang, Z.-G. Zhang, B.-T. Wang, J.-R. Zhang, and F.-W. Wang, Intrinsic ultralow lattice thermal conductivity in the full-Heusler compound Ba_2AgSb , *Phys. Rev. Applied* **17**, 034023 (2022).
- [109] M. Mohamed, B. Bouadjemi, M. Houari, A. Zitouni, T. Lantri, S. Haid, S. Bentata, B. Bouhaf, Z. Aziz, and R. Khenata, Electronic structure, mechanical and thermoelectric properties of the full-Heusler Ba_2AgZ ($\text{Z} = \text{Bi}, \text{Sb}$) alloys: insights from DFT study, *Indian J. Phys.* **95** (2021).

Paper IV

Discarded gems: Thermoelectric performance of materials with bandgap emerging at the hybrid-functional level

K. Berland, O. M. Løvvik, and R. Tranås

PUBLISHED. Appl. Phys. Lett., 119(8), 081902, (2021)

Discarded gems: Thermoelectric performance of materials with band gap emerging at the hybrid-functional level

Cite as: Appl. Phys. Lett. **119**, 081902 (2021); <https://doi.org/10.1063/5.0058685>

Submitted: 01 June 2021 . Accepted: 05 August 2021 . Published Online: 24 August 2021

 Kristian Berland,  Ole Martin Løvvik, and  Rasmus Tranås



View Online



Export Citation

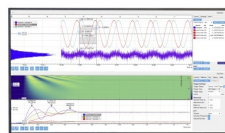


CrossMark



Challenge us.

What are your needs for periodic signal detection?



Zurich Instruments

Discarded gems: Thermoelectric performance of materials with band gap emerging at the hybrid-functional level

Cite as: Appl. Phys. Lett. **119**, 081902 (2021); doi: 10.1063/5.0058685

Submitted: 1 June 2021 · Accepted: 5 August 2021 ·

Published Online: 24 August 2021



View Online



Export Citation



CrossMark

Kristian Berland,^{1,a)}  Ole Martin Løvvik,^{2,3}  and Rasmus Tranås¹ 

AFFILIATIONS

¹Department of Mechanical Engineering and Technology Management, Norwegian University of Life Sciences, NO-1432 Ås, Norway

²SINTEF Industry, NO-0314 Oslo, Norway

³Centre for Materials Science and Nanotechnology, Department of Physics, University of Oslo, NO-0316 Oslo, Norway

Note: This paper is part of the APL Special Collection on Thermoelectric Materials Science and Technology Towards Applications.

a) Author to whom correspondence should be addressed: kristian.berland@nmbu.no

ABSTRACT

A finite electronic band gap is a standard filter in high-throughput screening of materials using density functional theory (DFT). However, because of the systematic underestimation of band gaps in standard DFT approximations, a number of compounds may be incorrectly predicted metallic. In a more accurate treatment, such materials may instead appear as low band gap materials and could have good thermoelectric properties if suitable doping is feasible. To explore this possibility, we performed hybrid functional calculations on 1093 cubic materials listed in the MATERIALS PROJECT database with four atoms in the primitive unit cell, spin-neutral ground state, and a formation energy within 0.3 eV of the convex hull. Out of these materials, we identified eight compounds for which a finite band gap emerges. Evaluating electronic and thermal transport properties of these compounds, we found the compositions MgSc₂Hg and Li₂CaSi to exhibit promising thermoelectric properties. These findings underline the potential of reassessing band gaps and band structures of compounds to identify additional potential thermoelectric materials.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0058685>

Thermoelectrics, with their ability to turn temperature gradients into electricity, can contribute to making the transition into a green economy with reduced greenhouse emission by recovering some of the waste heat generated in various industrial processes.^{1–3} While thermoelectric materials have traditionally not been sufficiently efficient for this task, great strides forward have been made in recent years. This has in turn intensified the hunt for novel thermoelectric materials,^{4–10} including the adoption of high-throughput screening and material informatics^{11,12} approaches.

The thermoelectric figure-of-merit $ZT = \sigma S^2 T / (\kappa_e + \kappa_l)$, which is the measure of the conversion efficacy, is given by the conductivity σ , the Seebeck coefficient S , the electronic κ_e and lattice thermal κ_l conductivity. Among these, all but κ_l are strongly linked to the electronic band structure. The electronic band gap E_{gap} is a particularly important parameter, determining the temperature for the onset of minority carrier transport, which causes a marked drop in S . It also has an indirect influence on the band curvature, i.e., as revealed by $\mathbf{k} \cdot \mathbf{p}$ -theory.¹³ Following Sofó and Mahan,¹⁴ a band gap of approximately 6–10 $k_B T$

has been traditionally considered attractive. However, their analysis was based on a direct band gap model with a single valley. Given its link to the band curvature, the band gap E_{gap} can also be viewed as a scale factor making a low band gap material more prone to exhibit multiple valleys in multipocketed band structures,¹⁵ nonetheless, the need to limit bipolar transport has made the existence of a finite band gap a standard criterion in most screening studies.⁷ Recently, attention has been broadened to other types of materials: Semi-metals with a strong asymmetry between conduction and valence bands have been marked as potential thermoelectric materials.^{16,17} Gapped metallic systems, which possess a band gap within the conduction or valence band, could also potentially exhibit good thermoelectric properties, once the band edge is sufficiently doped toward the Fermi level.¹⁸

A completely different reason for not discarding predicted metallic systems is that a number of them might have been mislabeled due to various approximations used in density functional theory (DFT).¹⁹ In particular, the commonly used generalized gradient approximation (GGA) systematically underestimates band gaps.^{20,21} This is less the

case for hybrid functionals,²² which mix a fraction of “exact” Fock exchange with the GGA.^{23,24} In the empirical linear relations between experimental and computed band gaps of Morales-García *et al.*,²¹ the offset of about 0.92 eV roughly indicates that compounds with a band gap smaller than this are likely to be incorrectly predicted as metallic by GGA.

In this work, we computed the band gap of 1093 cubic nonmagnetic materials listed in the MATERIALS PROJECT database²⁵ with four atoms in the primitive unit cell and a formation energy within 0.3 eV of the convex hull. These compounds include the full Heusler compounds with space group $Fm\bar{3}m$, inverse Heuslers with space group $F\bar{4}3m$ (both with composition X_2YZ), and binary AB_3 compounds. This reassessment resulted in eight compounds that were possibly mislabeled metallic by GGA. DFT calculations were performed using the VASP^{26–29} software package. The consistent-exchange van der Waals vdW-DF-cx functional^{30,31} was used for obtaining relaxed crystal structures and lattice thermal conductivities. While mostly used for modeling non-covalently bonded solids, recent studies have shown that vdW-DF-cx can improve structure and energetics compared to that of GGA of ionic and covalently bonded structures as well.^{32–34} To identify materials that could possess a band gap at the hybrid level, we first computed the band gap using merely a $4 \times 4 \times 4$ k -sampling of the Brillouin zone, including spin-orbit coupling using the HSE06^{23,35} hybrid functional. Such a low sampling can result in inaccurate Kohn-Sham energies, and we acknowledge that there is a slight risk that some compounds with very low band gap are missed. However, generally, the coarse sampling will cause a few systems to incorrectly appear with a finite or too large band gap. All systems with a finite band gap in the first stage were, therefore, reassessed with a $12 \times 12 \times 12$ k -point sampling of the Fock operator and charge density, which is used to compute the band structure path using 101 k -points along $W-L-\Gamma-X-K$ to obtain an accurate band gap. For the new band gap compounds, the electronic transport properties were computed with the Boltzmann transport equation in the constant relaxation time approximation with $\tau = 10^{-14}$ s using BOLTZTRAP.³⁶ To ensure dense grid sampling, we used a corrected $k \cdot p$ -based interpolation method,^{37,38} using the same computational parameters as in Ref. 39. The lattice thermal conductivity, κ_ℓ , was computed using the temperature-dependent effective potential (TDEP) method.^{40,41} A canonical ensemble was used to generate 50 uncorrelated configurations based on a $3 \times 3 \times 3$ repetition of the relaxed primitive cell.⁴² The positions and forces of the supercells allowed for extraction of second- and third-order force constants. The cutoff for second-order interactions was set to 7 Å, while for third-order, a cutoff slightly larger than half the width of the supercell was used. Reciprocal space discretization for Brillouin zone integrations was done using a $35 \times 35 \times 35$ q -point grid. Isotope scattering was also included. All supplementary GGA calculations in this Letter were based on the version of Perdew–Burke–Ernzerhof (PBE).⁴³

Among the 1093 compounds materials examined, eight compounds have a band gap at the HSE06 level as listed in Table I; corresponding band structures are provided in the supplementary material (SM).

Figure 1 shows the computed κ_ℓ for the identified compounds. Very low values of κ_ℓ were found for Ba_2HgPb ranging from 0.46 W/mK at 300 K to 0.17 W/mK at 800 K. This compound was also studied by He *et al.*⁴⁴ predicting values of κ_ℓ somewhat larger than

TABLE I. Properties of new band gap compounds.

Compound	No. of valence	E_{hull} (MP)	Band gap (eV)
AlVFe ₂	24	0	0.78
Ba ₂ HgPb	20	0	0.06
HfSnRu ₂	24	0	0.21
Li ₂ CaSi	10	0	0.01
MgSc ₂ Hg	20	0	0.23
TaInRu ₂	24	0	0.05
TiSiO ₂	24	0	0.55
VGaFe ₂	24	0	0.66

ours. Possible reasons for this difference include their use of a compressive sensing lattice dynamic technique⁴⁵ to obtain third-order force constants and other technical details, differing exchange correlation functionals, and the phonon-mode renormalization inherit to TDEP. Comparing TDEP and PHONOPY, Feng *et al.*⁴⁶ found lower κ_ℓ for TDEP than with the standard-finite difference approach and argued that TDEP is better suited to describe low- κ_ℓ materials.

Based solely on Fig. 1, only Ba_2HgPb , Li_2CaSi and $MgSc_2Hg$ have low enough κ_ℓ to conceivably be good thermoelectric materials. Yet, the literature is riddled with examples of how various disorder-related scattering mechanisms, such as grain boundaries, defects, and substitutions, can dramatically lower κ_ℓ .^{47–54} For this reason, we used $\kappa_\ell = 4$ W/mK as the maximum for all materials in further comparisons.

Figure 2 plots the optimal doping concentration against peak ZT for each of the compounds in temperature steps of 100 K from 300 to 800 K, for doping concentration between 10^{19} cm^{-3} and $3 \times 10^{21} \text{ cm}^{-3}$. Based on this plot, we deem Li_2CaSi and $MgSc_2Hg$ to have great potential as thermoelectric n -type materials, while $MgSc_2Hg$ and $AlVFe_2$ have some potential as p -type thermoelectrics. n -type $AlVFe_2$ has been studied earlier theoretically at the hybrid functional level⁵⁵ and experimentally.^{56,57} The study of Mikami *et al.*⁵⁶ measured ZT in a similar range once doping and sublattice disorder were introduced. While Li_2CaSi is reported as stable in the $Fm\bar{3}m$ Heusler phase in MATERIALS PROJECT;

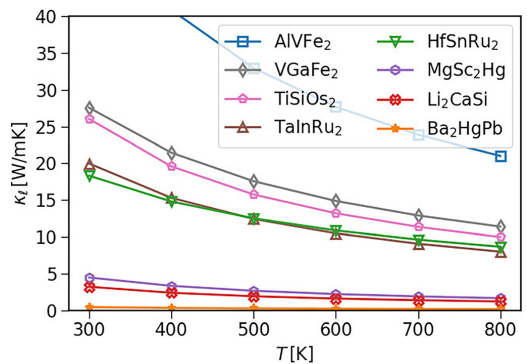


FIG. 1. Lattice thermal conductivity of identified compounds computed with TDEP.

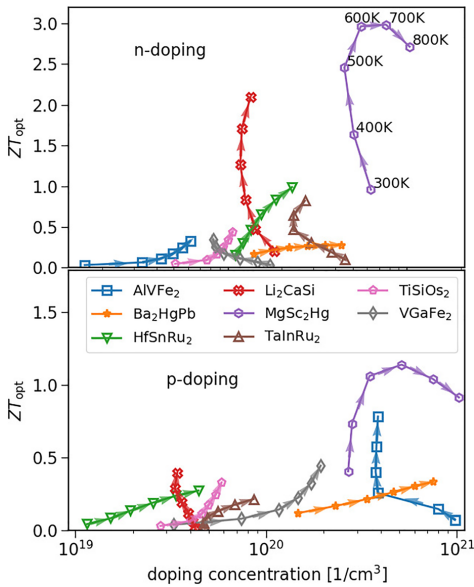


FIG. 2. Optimized ZT at different temperatures from 300 to 800 K with arrows indicating increasing temperature in steps of 100 K. The vertical axis gives the optimized ZT , while the horizontal gives the corresponding doping concentration.

experimentally, it has been crystallized in the orthorhombic $Pmnm$ phase.⁵⁸ The related Li_2CaSn , on the other hand, does crystallize in the Heusler phase. No experimental realizations of $MgSc_2Hg$ are known to us.

The origin of the high ZT of Li_2CaSi and $MgSc_2Hg$ can be related to their band structures as shown in Fig. 3. The band structure of Li_2CaSi exhibits some noticeable features: (i) Dirac points at the Γ -point with a band opening of 0.01 eV, (ii) near convergence of a number of additional bands at the Γ -point, (iii) electron bands that are flat in the Γ - X direction, but dispersive in the X - K direction. In our study, we find similar features in the band structure of $HfSnRu_2$ and

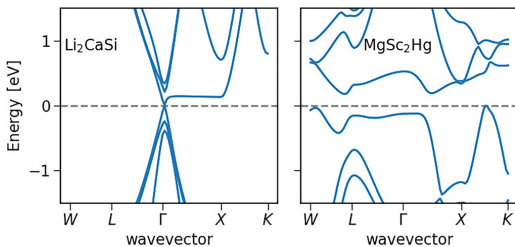


FIG. 3. Electronic band structures of Li_2CaSi and $MgSc_2Hg$.

$TaInRu_2$, which also exhibit relatively high ZT for n -type doping. While Bilc *et al.*⁵⁵ argued that band structures of this type can give rise to high ZT due to their effectively low-dimensional transport, Park *et al.*⁵⁹ demonstrated that flat-and-dispersive band structures, specifically for the case of Fe_2TiSi , can cause large effective scattering phase-space, which significantly reduces the power factor. In contrast, $MgSc_2Hg$ band structure has a multi-valley structure, in particular, in the conduction band. In fact, with the exception of the highly dispersive band in the X -point, the band structure can be viewed as a partial realization of a δ -function like transport distribution function, which in the analysis of Mahan and Sofo is optimal for thermoelectric performance.⁶⁰ Other cubic structures, such as the 10-valence electron full-Heusler compounds predicted by He *et al.*,⁶¹ also have similarly attractive band structure features.

Figure 4(a) shows Pisarenko-type plots for the thermoelectric properties of $MgSc_2Hg$ at 300, 600, and 800 K, while Fig. 4(b) shows the underlying distribution functions giving rise to these properties. They are related through the transport distribution function $\Sigma(\epsilon)$ as follows:^{56,60}

$$\sigma = e^2 \int d\epsilon \Sigma(\epsilon - \mu_F) f_1(\epsilon - \mu_F), \quad (1)$$

$$\sigma S = (e/T) \int d\epsilon (\epsilon - \mu_F) \Sigma(\epsilon - \mu_F) f_1(\epsilon - \mu_F), \quad (2)$$

$$\kappa_0 = (1/T) \int d\epsilon \Sigma(\epsilon - \mu_F) (\epsilon - \mu_F)^2 f_1(\epsilon - \mu_F), \quad (3)$$

where μ_F is the Fermi level and f_1 is the Fermi window, given by the derivative of the Fermi-Dirac function, $f_1(\epsilon - \mu_F) = -df_{FD}/d\epsilon$. The open-circuit electronic thermal conductivity κ_0 is related to the closed-circuit by $\kappa_e = \kappa_0 - T\sigma S^2$. The temperature dependence stems explicitly from the Fermi-Dirac function and implicitly from the temperature dependence of μ_F . A dashed line indicates the peak of $\Sigma(\epsilon)$ for comparison with the band structure as shown in Fig. 3. Figures show that for $MgSc_2Hg$, the magnitude of κ_e is a key factor limiting ZT at elevated temperatures. They also show that a minimum in κ_e at 600 and 800 K occurs at a higher doping concentration than what maximizes S . They both reach extreme values due to a minimum in the bipolar transport, but the second moment $(\epsilon - \mu_F)^2$ entering into κ_0 [Eq. (3)] shifts the optimum of κ_e to a higher doping concentration. The figure also indicates that the rapidly rising $\Sigma(\epsilon)$ up to the peak occurring at 0.55 eV explains why S can be quite large despite a low band gap even at high doping concentrations. At the same time, it shows that this rapid rise is the cause of the large values of κ_e at high temperatures.

Figures 4(c) and 4(d) show corresponding results for Li_2CaSi . It is interesting to note that while the band gap is tiny, the low $\Sigma(\epsilon)$ in the valence band makes this compound resemble a wide band-gap semiconductor. In fact, at optimal doping concentration, the bipolar transport almost entirely occurs within the conduction band. While the limited bipolar transport results in higher S at lower doping concentrations, Li_2CaSi lacks the beneficial peak in $\Sigma(\epsilon)$ present in $MgSc_2Hg$, which limits κ_e at higher temperatures and doping concentrations. The low band gap of Li_2CaSi makes it interesting to also consider the properties of Li_2CaSi as predicted at the GGA level. In this case, a finite gap is retained at the Γ point but the material is self-doped and the flat-and-dispersive band crosses the Fermi level at zero

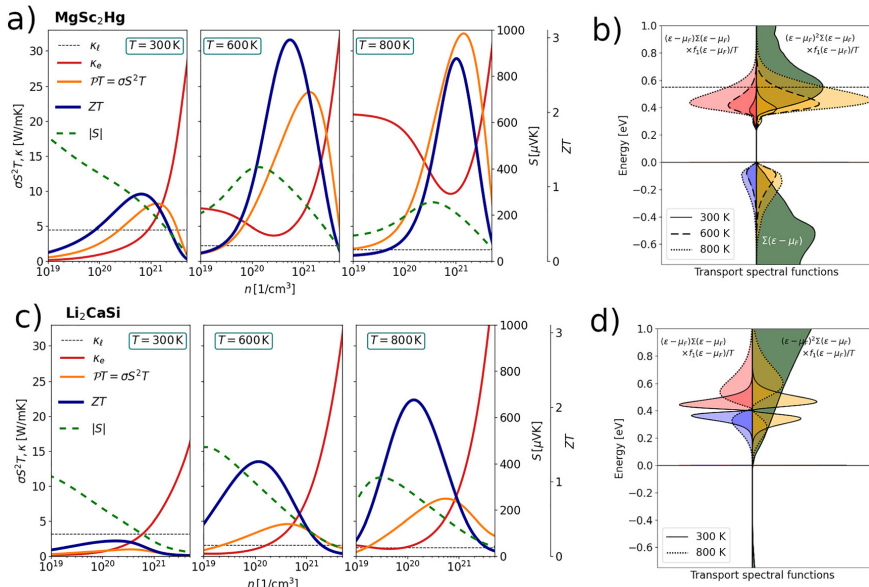


FIG. 4. (a) and (c) Thermoelectric properties of MgSc_2Hg [Li_2CaSi] as a function of doping concentration at 300, 600, and 800 K. In (b) and (d), the green background shows the corresponding transport distribution function $\Sigma(\epsilon)$. The left and right sides of the vertical axis show the contributions to the first and second moments of the $\Sigma(\epsilon)$ weighted by the derivative $f_1(\epsilon) = -df_{\text{FD}}(\epsilon)/d\epsilon$, which is proportional to, respectively, $T\sigma S$ and the closed circuit thermal conductivity κ_0 . Results for 600 K omitted for clarity in (d).

extrinsic doping. An optimal $ZT = 0.76$ at 800 K is predicted—further details in the [supplementary material](#).

While we in this study assessed the properties of 1093 four-atom materials using sub-converged hybrid functional calculations, other approaches could also be worth exploring. We investigated the potential of analyzing the GGA-level density of states, in which “narrowing” could hint of a finite band gap. Details can be found in the [supplementary material](#). Interestingly, this approach clearly indicated all compounds except the MgSc_2Hg compound; precisely, the property that made this material into a promising thermoelectric, i.e., the high density of states close to the band edges at the hybrid level, made the density-of-states narrowing at the GGA-level vanish. We, therefore, do not generally recommend this approach to uncover high performance thermoelectric materials.

In this Letter, we have demonstrated that the use of GGA-level band structures can cause promising thermoelectric materials to be discarded because they are falsely predicted to be metallic. This was illustrated with the finding of new thermoelectric compounds with a band gap appearing at the hybrid functional level: Out of the 1093 studied compounds, eight were identified with a band gap by hybrid calculations and not by GGA calculations. Out of these, a few were also promising for thermoelectric applications: MgSc_2Hg , Li_2CaSi , and to some extent AlVFe_2 . The Heusler MgSc_2Hg compound, in particular, exhibits excellent potential as a thermoelectric material. We are not aware of any experimental realization of this compound or

in-depth stability analysis. Moreover, the toxicity of Hg reduces the attractiveness of this compound for general-purpose applications. In addition to realizability, we stress that the use of a constant relaxation-time approximation is a coarse approximation. The inclusion of proper electron-phonon scattering can have a decisive impact upon the power factor and predicted ZT properties.⁶² Another concern is whether hybrid functionals in fact do provide accurate band structures for these intermetallic compounds, which can be investigated for instance by performing *GW*-level calculations, as earlier done for selected half Heuslers.⁶³ Despite these caveats, our study clearly underlines that high performing thermoelectric materials can be uncovered through the reassessment of electronic band gaps.

On a final note, it is interesting that three compounds with lowest κ_f and two of the compounds with the highest *n*-type ZT violated the octet rule or the corresponding 18- and 24-electron rules. This violation is a feature shared with the well-known thermoelectric PbTe and related compounds.^{61,64} The existence of lone *s*-pairs has earlier been linked to low thermal conductivity.^{61,65,66} One could speculate that going beyond GGA could be particularly pertinent for the electronic band structure of octet violating systems, similar to what we found earlier for PbTe .³⁸

See the [supplementary material](#) for computed band structures at the HSE06 level, density of states at the PBE level of theory, and band structure and *n*-type thermoelectric properties of Li_2CaSi .

The computations were performed on resources provided by UNINETT Sigma2—the National Infrastructure for High Performance Computing and Data Storage in Norway. This work was in part funded by the Allotherm project (Project No. 314778) supported by the Research Council of Norway.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- D. M. Rowe and G. Min, "Evaluation of thermoelectric modules for power generation," *J. Power Sources* **73**, 193–198 (1988).
- M. Dresselhaus, "Overview of thermoelectrics for thermal to electrical energy conversion," *AIP Conf. Proc.* **1519**, 36–39 (2013).
- G. D. Mahan, "Introduction to thermoelectrics," *APL Mater.* **4**, 104806 (2016).
- S. Bhattacharya and G. K. H. Madsen, "High-throughput exploration of alloying as design strategy for thermoelectrics," *Phys. Rev. B* **92**, 085205 (2015).
- S. Bhattacharya and G. K. H. Madsen, "A novel p-type half-Heusler from high-throughput transport and defect calculations," *J. Mater. Chem. C* **4**, 11261–11268 (2016).
- W. Chen, J.-H. Pöhl, G. Hautier, D. Broberg, S. Bajaj, U. Aydemir, Z. M. Gibbs, H. Zhu, M. Asta, G. J. Snyder, B. Meredig, M. A. White, K. Persson, and A. Jain, "Understanding thermoelectric properties from high-throughput calculations: Trends, insights, and comparisons with experiment," *J. Mater. Chem. C* **4**, 4414–4426 (2016).
- R. Li, X. Li, L. Xi, J. Yang, D. J. Singh, and W. Zhang, "High-throughput screening for advanced thermoelectric materials: Diamond-like ABX₂ compounds," *ACS Appl. Mater. Interfaces* **11**, 24859 (2019).
- C. Barreteau, J.-C. Crivello, J.-M. Joubert, and E. Alleno, "Looking for new thermoelectric materials among TMX intermetallics using high-throughput calculations," *Comput. Mater. Sci.* **156**, 96–103 (2019).
- L. Chen, H. Tran, R. Batra, C. Kim, and R. Ramprasad, "Machine learning models for the lattice thermal conductivity prediction of inorganic materials," *Comput. Mater. Sci.* **170**, 109155 (2019).
- Y. Iwasaki, I. Takeuchi, V. Stanev, A. G. Kusne, M. Ishida, A. Kirihara, K. Ihara, R. Sawada, K. Terashima, H. Someya, K-i Uchida, E. Saitoh, and S. Yorozu, "Machine-learning guided discovery of a new thermoelectric material," *Sci. Rep.* **9**, 2751 (2019).
- K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature* **559**, 547 (2018).
- J. Hill, G. Mulholland, K. Persson, R. Seshadri, C. Wolverton, and B. Meredig, "Materials science with large-scale data and informatics: Unlocking new opportunities," *MRS Bull.* **41**, 399–409 (2016).
- J. M. Luttinger, "Quantum theory of cyclotron resonance in semiconductors: general theory," *Phys. Rev.* **102**, 1030–1041 (1956).
- O. Sofo and G. D. Mahan, "Optimum band gap of a thermoelectric material," *Phys. Rev. B* **49**, 4565–4570 (1994).
- N. Wang, M. Li, H. Xiao, Z. Gao, Z. Liu, X. Zu, S. Li, and L. Qiao, "Band degeneracy enhanced thermoelectric performance in layered oxyselenides by first-principles calculations," *npj Comput. Mater.* **7**, 18 (2021).
- M. Markov, X. Hu, H.-C. Liu, N. Liu, S. J. Poon, K. Esfarjani, and M. Zebarjadi, "Semi-metals as potential thermoelectric materials," *Sci. Rep.* **8**, 9876 (2018).
- M. Markov, S. E. Rezaei, S. N. Sadeghi, K. Esfarjani, and M. Zebarjadi, "Thermoelectric properties of semimetals," *Phys. Rev. Mater.* **3**, 095401 (2019).
- F. Ricci, A. Dunn, A. Jain, G.-M. Rignanesse, and G. Hautier, "Gapped metals as thermoelectric materials revealed by high-throughput screening," *J. Mater. Chem. A* **8**, 17579–17594 (2020).
- O. I. Malyi and A. Zunger, "False metals, real insulators, and degenerate gapped metals," *Appl. Phys. Rev.* **7**, 041310 (2020).
- A. Pribram-Jones, D. A. Gross, and K. Burke, "DFT: A Theory Full of Holes," *Annu. Rev. Phys. Chem.* **66**, 283–304 (2015).
- A. Morales-García, R. Valero, and F. Illas, "An empirical, yet practical way to predict the band gap in solids by using density functional band structure calculations," *J. Phys. Chem. C* **121**, 18862–18866 (2017).
- S. Kim, M. Lee, C. Hong, Y. Yoon, H. An, D. Lee, W. Jeong, D. Yoo, Y. Kang, Y. Youn, and S. Han, "A band-gap database for semiconducting inorganic materials calculated with hybrid functional," *Sci. Data* **7**, 387 (2020).
- J. P. Perdew and M. Ernzerhof, "Rationale for mixing exact exchange with density functional approximations," *J. Chem. Phys.* **105**, 9982 (1996).
- J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential," *J. Chem. Phys.* **118**, 8207–8215 (2003).
- 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, "Commentary: The materials project: A materials genome approach to accelerating materials innovation," *APL Mater.* **1**, 011002 (2013).
- G. Kresse and J. Hafner, "Ab Initio molecular dynamics for liquid metals," *Phys. Rev. B* **47**, 558–561 (1993).
- G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," *Comput. Mater. Sci.* **6**, 15–50 (1996).
- G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Phys. Rev. B* **54**, 11169–11186 (1996).
- M. Gajdos, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, "Linear optical properties in the projector-augmented wave methodology," *Phys. Rev. B* **73**, 045112 (2006).
- K. Berland and P. Hyldgaard, "Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional," *Phys. Rev. B* **89**, 035412 (2014).
- K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist, "Van der Waals forces in density functional theory: A review of the vdW-DF method," *Rep. Prog. Phys.* **78**, 066501 (2015).
- D. O. Lindroth and P. Erhart, "Thermal transport in van der Waals solids from first-principles calculations," *Phys. Rev. B* **94**, 115205 (2016).
- L. Gharaee, P. Erhart, and P. Hyldgaard, "Finite-temperature properties of nonmagnetic transition metals: Comparison of the performance of constraint-based semilocal and nonlocal functionals," *Phys. Rev. B* **95**, 085147 (2017).
- P. Hyldgaard, Y. Jiao, and V. Shukla, "Screening nature of the van der Waals density functional method: A review and analysis of the many-body physics foundation," *J. Phys. Condens. Matter* **32**, 393001 (2020).
- A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the exchange screening parameter on the performance of screened hybrid functionals," *J. Chem. Phys.* **125**, 224106 (2006).
- G. K. Madsen and D. J. Singh, "BoltzTraP. A code for calculating band-structure dependent quantities," *Comput. Phys. Commun.* **175**, 67–71 (2006).
- K. Berland and C. Persson, "Enabling accurate first-principle calculations of electronic properties with a corrected $k \cdot p$ scheme," *Comput. Mater. Sci.* **134**, 17–24 (2017).
- K. Berland and C. Persson, "Thermoelectric transport of GaAs, InP, and PbTe: Hybrid functional with $k \cdot p$ interpolation versus scissor-corrected generalized gradient approximation," *J. Appl. Phys.* **123**, 205703 (2018).
- K. Berland, N. Shulumba, O. Hellman, C. Persson, and O. M. Løvvik, "Thermoelectric transport trends in group 4 half-Heusler alloys," *J. Appl. Phys.* **126**, 145102 (2019).
- O. Hellman, I. A. Abrikosov, and S. I. Simak, "Lattice dynamics of anharmonic solids from first principles," *Phys. Rev. B* **84**, 180301 (2011).
- O. Hellman and D. A. Broido, "Phonon thermal transport in Bi_2Te_3 from first principles," *Phys. Rev. B* **90**, 134309 (2014).
- N. Shulumba, O. Hellman, and A. J. Minnich, "Lattice thermal conductivity of polyethylene molecular crystals from first-principles including nuclear quantum effects," *Phys. Rev. Lett.* **119**, 185901 (2017).
- J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- J. He, M. Amsler, Y. Xia, S. S. Naghavi, V. I. Hegde, S. Hao, S. Goedecker, V. Ozoliņš, and C. Wolverton, "Ultralow thermal conductivity in full Heusler semiconductors," *Phys. Rev. Lett.* **117**, 046602 (2016).
- F. Zhou, W. Nielson, Y. Xia, and V. Ozoliņš, "Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations," *Phys. Rev. Lett.* **113**, 185501 (2014).

- ⁴⁶Z. Feng, Y. Fu, Y. Zhang, and D. J. Singh, "Characterization of rattling in relation to thermal conductivity: Ordered half-Heusler semiconductors," *Phys. Rev. B* **101**, 064301 (2020).
- ⁴⁷B. Abeles, "Lattice thermal conductivity of disordered semiconductor alloys at high temperatures," *Phys. Rev.* **131**, 1906–1911 (1963).
- ⁴⁸Z. Tian, J. Garg, K. Esfarjani, T. Shiga, J. Shiomi, and G. Chen, "Phonon conduction in PbSe, PbTe, and PbTe_{1-x}Se_x from first-principles calculations," *Phys. Rev. B* **85**, 184303 (2012).
- ⁴⁹A. Katre, J. Carrete, and N. Mingo, "Unraveling the dominant phonon scattering mechanism in the thermoelectric compound ZrNiSn," *J. Mater. Chem. A* **4**, 15940 (2016).
- ⁵⁰M. Arrigoni, J. Carrete, N. Mingo, and G. K. H. Madsen, "First-principles quantitative prediction of the lattice thermal conductivity in random semiconductor alloys: The role of force-constant disorder," *Phys. Rev. B* **98**, 115205 (2018).
- ⁵¹J. Carrete, N. Mingo, S. Wang, and S. Curtarolo, "Nanograined half-Heusler semiconductors as advanced thermoelectrics: An ab initio high-throughput statistical study," *Adv. Funct. Mater.* **24**, 7427–7432 (2014).
- ⁵²W. Li, L. Lindsay, D. A. Broido, D. A. Stewart, and N. Mingo, "Thermal conductivity of bulk and nanowire Mg₂Si_xSn_{1-x} alloys from first principles," *Phys. Rev. B* **86**, 174307 (2012).
- ⁵³S. N. H. Eliassen, A. Katre, G. K. H. Madsen, C. Persson, O. M. Løvvik, and K. Berland, "Lattice thermal conductivity of Ti_xZr_yHf_{1-x-y}NiSn half-Heusler alloys calculated from first principles: Key role of nature of phonon modes," *Phys. Rev. B* **95**, 045202 (2017).
- ⁵⁴M. Schrade, K. Berland, S. N. H. Eliassen, M. N. Guzik, C. Echevarria-Bonet, M. H. Sørbj, P. Jenus, B. C. Hauback, R. Tofan, A. E. Gunnæs, C. Persson, O. M. Løvvik, and T. G. Finstad, "The role of grain boundary scattering in reducing the thermal conductivity of polycrystalline XNiSn (X = Hf, Zr, Ti) half-Heusler alloys," *Sci. Rep.* **7**, 13760 (2017).
- ⁵⁵D. I. Bilc, G. Hautier, D. Waroquiers, G.-M. Rignanes, and P. Ghosez, "Low-dimensional transport and large thermoelectric power factors in bulk semiconductors by band engineering of highly directional electronic states," *Phys. Rev. Lett.* **114**, 136601 (2015).
- ⁵⁶M. Mikami, Y. Kinemuchi, K. Ozaki, Y. Terazawa, and T. Takeuchi, "Thermoelectric properties of tungsten-substituted Heusler Fe₂VAl alloy," *J. Appl. Phys.* **111**, 093710 (2012).
- ⁵⁷M. Vasundhara, V. Srinivas, and V. V. Rao, "Electronic transport in Heusler-type Fe₂VAl_{1-x}M_x alloys (M = B, In, Si)," *Phys. Rev. B* **77**, 224415 (2008).
- ⁵⁸D. Stoiber, M. Bobnar, P. Höhn, and R. Niewa, "Lithium alkaline earth tetrelides of the type Li₂AeTt (Ae = Ca, Ba, Tt = Si, Ge, Sn, Pb): Synthesis, crystal structures and physical properties," *Z. Naturforsch. B* **72**, 847–853 (2017).
- ⁵⁹J. Park, Y. Xia, and V. Ozoliņš, "High thermoelectric power factor and efficiency from a highly dispersive band in Ba₂BiAu," *Phys. Rev. Appl.* **11**, 014058 (2019).
- ⁶⁰G. D. Mahan and J. O. Sofo, "The best thermoelectric," *Proc. Natl. Acad. Sci.* **93**, 7436–7439 (1996).
- ⁶¹J. He, Y. Xia, S. S. Naghavi, V. Ozoliņš, and C. Wolverton, "Designing chemical analogs to PbTe with intrinsic high band degeneracy and low lattice thermal conductivity," *Nat. Commun.* **10**, 719 (2019).
- ⁶²J. Zhou, H. Zhu, T.-H. Liu, Q. Song, R. He, J. Mao, Z. Liu, W. Ren, B. Liao, D. J. Singh, Z. Ren, and G. Chen, "Large thermoelectric power factor from crystal symmetry-protected non-bonding orbital in half-Heuslers," *Nat. Commun.* **9**, 1721 (2018).
- ⁶³M. Zahedifar and P. Kratzer, "Band structure and thermoelectric properties of half-Heusler semiconductors from many-body perturbation theory," *Phys. Rev. B* **97**, 035204 (2018).
- ⁶⁴J. R. Sootsman, R. J. Pcionek, H. Kong, C. Uher, and M. G. Kanatzidis, "Strong reduction of thermal conductivity in nanostructured PbTe prepared by matrix encapsulation," *Chem. Mater.* **18**, 4993–4995 (2006).
- ⁶⁵O. Delaire, J. Ma, K. Marty, A. F. May, M. A. McGuire, M.-H. Du, D. J. Singh, A. Podlesnyak, G. Ehlers, M. D. Lumsden, and B. C. Sales, "Giant anharmonic phonon scattering in PbTe," *Nat. Mater.* **10**, 614–619 (2011).
- ⁶⁶M. K. Jana, K. Pal, U. V. Waghmare, and K. Biswas, "The origin of ultralow thermal conductivity in InTe: Lone-pair-induced anharmonic rattling," *Angew. Chem. Int. Ed* **55**, 7792–7796 (2016).

ISBN: 978-82-575-2042-7

ISSN: 1894-6402



Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no