Atomic-Scale Simulations of Si-Ge-Sn Alloys Using Deep-Learning-Based Interatomic Potentials

by

Anirudh Dabir

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Science

Approved April 2021 by the
Graduate Supervisory Committee:

Houlong Zhuang, Chair
Yang Jiao
Qiong Nian

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

Accurate knowledge and understanding of thermal conductivity is very important in a wide variety of applications both at microscopic and macroscopic scales. Estimation, however varies widely with respect to scale and application. At a lattice level, calculation of thermal conductivity of any particular alloy require very heavy computation even for a relatively small number of atoms. This thesis aims to run conventional molecular dynamic simulations for a particular supercell and then employ a machine learning based approach and compare the two in hopes of developing a method to greatly reduce computational costs as well as increase the scale and time frame of these systems. Conventional simulations were run using interatomic potentials based on density function theory-based *ab initio* calculations. Then deep learning neural network based interatomic potentials were used run similar simulations to compare the two approaches.

ACKNOWLEDGEMENTS

I want to take this opportunity to thank my professor and mentor, **Dr. Houlong Zhuang**, Assistant Professor, School for Engineering of Matter, Transport and Energy, Arizona State University for his constant support and guidance. His support and dedication has played a large part in me completing my thesis on time. His advice, guidance and scientific approach will always stay with my in my endeavours in the future.

I owe a debt of gratitude to Dr. Duo Wang, a former member of Dr. Zhuang's team. His support and research work has played a vital role in the core work of my thesis. His timely support and guidance have been invaluable.

I wish to thank Dr. Lei Liu, another former member of Dr. Zhuang's team. His role in helping setting up environments and software packages on the cluster computing services enabled me to start my work in a timely manner.

I wish to thank the research and computing admins maintaining the Agave Cluster Computing Services at Arizona State Universities for timely technical support and vigilant maintenance of the cluster, ensuring adequate services to all those who needed high performance computing for projects such as mine.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Thermal Conductivity and Its Importance

A multitude of situations and applications require prior knowledge of the thermal conductivity of a particular substance either as a bulk material or at an atomic unit cell level. Choosing the right material for the application at hand quite often will depend on that material possessing the adequate value of thermal conductivity. For instance, hot food from a restaurant is often packed in plastic or styrofoam boxes since these materials prevent the loss of heat to the outside environment, thus keeping the contents hotter for longer. These materials are thermal insulators, in that they hinder the transport of heat through them. This is a direct result of low thermal conductivity. On the other hand, metallic vessels and containers used in kitchens are thermal conductors. Meaning, they readily allow the transport of heat through them because metals in general have high thermal conductivity.

1.2 Thermal Conductivity of Bulk Materials

Thermal conductivity of bulk materials refers to the ability of a material to transport heat through its volume at a macroscopic scale. The interactions that take place between individual atoms at a unit cell level are irrelevant here. The scenario of the restaurant in the previous paragraph is a perfect example of such a case. Applications like temperature measurement and active cooling of certain systems also require good understanding of thermal conductivity. These scenarios use the thermoelectric effect to measure temperature or even create localized hot and cold zones.

1.2.1 The Thermoelectric Effect

The term "thermoelectric effect" is used to refer to three different phenomena – the Seebeck effect, the Peltier effect and the Thomson effect. Of the three, the first to be discovered was the Seebeck effect by Thomas Seebeck in 1821 [7]. He discovered that a temperature difference across a junction made from dissimilar metals resulted in an electric potential difference across that junction. As a result, the term "thermoelectric effect" commonly refers to the Seebeck effect.

The most common application of the Seebeck effect is to measure temperature. This is done by the use of a thermocouple. These devices find themselves to be useful anywhere starting from furnaces to fog machines. Being simple in principle and compact in design, thermocouples have a wide range of operating temperatures (-270 – 3000 C)[1]. In addition to this most thermocouples do not require expensive and/or rare materials as raw materials.

In 1834, the converse of the Seebeck effect was proven by Jean Peltier [7]. Known as the Peltier effect, here a temperature difference is created by applying an electric potential difference across a junction. This is the main operating principle behind devices like peltier coolers. These are devices that consume electric power in order to generate localized regions of high and low temperatures.

Both the Seebeck and Peltier effects have associated coefficients, which are properties of the materials that make up the junction. These coefficients dictate the voltage and temperature difference respectively and are assumed in most cases to be constants. However, in a significant number of cases, these coefficients are found to be varying with temperature, thus creating a gradient. This was identified by Lord Kelvin (William Thomson) in 1851 [1]. The Thomson effect considers the effects like joule heating into effect.

The Seebeck Effect

The Seebeck effect describes a phenomenon where a temperature difference across a junction results in an electric potential difference (voltage) across that junction. Devices like the thermocouple make use of this effect to measure temperature. The voltage developed can be estimated as follows

$$J = \sigma(-\nabla V + E_{emf})$$

where S is the Seebeck coefficient. Then, the current density can be estimated using the equation below.

$$E_{emf} = -S\nabla T$$

Where,

J : Current Density

σ : Local thermal conductivity

∇V : Local voltage

S : Seebeck coefficient

∇T : Temperature gradient

Measurement of temperature is not the only mainstream application of the Seebeck effect. Another class of devices that use this phenomenon are thermoelectric generators (TEGs). These devices use the heat flux from any source as an input and then output usable electrical energy. TEGs have no moving parts or are solid state, making them virtually maintenance free. This enables their use in unmanned/high stress environments where frequent servicing is either not possible or viable. TEGs have been used in applications where an already established process generates waste

heat, like in power plants and the automobile sector.

Waste heat from fossil fuel power plants can be used to generate usable electricity, thus increasing the efficiency of the plants. In the automobile sector, the heat generated during braking is used to power TEGs. The electricity generated can be used to power the vehicle's systems. And since this acts as an energy drain for the heat generated, braking efficiency is slightly increased as well.

Applications of TEGs need not be limited to recycling waste heat. There are cases where TEGs are the primary source of power as well. The heat generated by a radioactive isotope can be the heat source for TEGs that supply power to systems that cannot have bulky energy solutions. This is the idea behind the electricity generators on board multiple space probes. In a place where quite literally every gram of weight counts, TEGs offer a lighter and more reliable alternative to chemical energy storage.

The Peltier Effect

While the Seebeck effect generates a voltage from a temperature difference, a peltier device seeks to generate a temperature difference by consuming electric energy. Literally, this is the converse of the Seebeck effect. And like the Seebeck effect, there are range of devices that use this phenomenon and are applied in a wide variety of cases.

The most common application of the peltier effect is in peltier coolers, also known as thermoelectric coolers (TECs). These solid state devices can be considered to be the counterparts for TEGs. And given that the main outcome is the creation of localized regions of high and low temperature, these devices act as heat pumps. Given their lack of moving parts and small form factor, TEGs offer a compelling alternative to conventional cooling / heating systems in certain applications.

TEGs do not create a specific temperature for a given set of conductors and input current. Rather, they create a specific temperature difference. Using this effect,

the cooling of the "hot" side of the cooler can result in much lower than normal temperatures in the "cold" side. One of the best applications of this idea can be seen in the cloud chamber particle detector created by electrical engineer and popular YouTuber, Mehdi Sadagdar on his YouTube channel, ElectroBOOM [9].

In the video, Mr. Sadagdar creates a sealed glass chamber whose base is a very cold and dry surface. The air in in the chamber is evacuated and filled with alcohol vapors. Left undisturbed, any particles of radiation passing through the chamber collides with the alcohol vapour, creating a small but visible condensed cloud of alcohol, similar to a vapor cone around a fighter plane breaking the sound barrier. Usually, the cold surface is created using dry ice or even liquid nitrogen. Mr. Sadagdar goes another way, by using a bank of peltier coolers. He freezes the hot side in salt water, thus enabling the cold side to achieve temperatures cold enough to sustain alcohol vapors.

Peltier coolers can be theoretically used to cool electronic components as well. As part of my course work, I was involved in a project which explored the possibility of a TEC replacing a conventional heat sink and fan cooling solution for a low powered laptop CPU. The scope of the project was quite limited, and constraints were high. It was understood during this project that the thermal conductivity of the junction was important to take into account the joule heating effect. The junctions in the cooler had to be chosen in such a way that there would be minimal heat flow between the hot and cold sides of the cooler.

1.3 Lattice Thermal Conductivity

Study of thermal conductivity at a lattice level has its own importance. The study at a macroscopic level mainly focuses on the applications in real life. While the study at a microscopic or lattice level explores the physics behind it all. In most metals and semiconductors like silicon, lattice thermal conductivity is the dominant mode

of thermal transfer within material [10]. Hence, the study of lattice thermal conductivity provides insight into the mechanism by which heat is transported through the medium. In the case of semiconductors, lattice thermal conductivity is particularly useful in the electronics industry. For instance, doping of semiconductors to make particular types of diodes, junctions, etc. While the primary reactions for doping would be to alter properties like electrical resistance and diode characteristics, thermal conductivity becomes a very relevant factor when considering effects like thermal dissipation and efficiency. The subsequent chapters focus on the mechanisms behind thermal conduction and estimation of lattice thermal conductivity of a supercell. The structure is a medium entropy alloy, made out of silicon, germanium and tin. The concentration of tin will be varied and the value of thermal conductivity estimated.

Chapter 2

INTERATOMIC INTERACTIONS

In crystal lattice structures, the lattice vibrations themselves are a major contributing factor for the thermal conductivity of the unit cell[3]. These vibrations, or phonons are responsible for the rate of flow of free electrons within the unit cell, that in the end determine the the thermal conductivity of that cell. However, estimating the phonon interactions in a unit cell, even with a simple lattice structure is no easy task. There exists no unified theory that fits all cases; some provide high accuracy at the cost of system scaling and others vice-versa.

In the early stages of development of the field, popular theories included that Boltzmann kinetic approach and molecular dynamic simulations. However, given the limitations of computational technology of the era, these approaches had their drawbacks with the computational costs being high and lacking in accuracy in many cases. Later advancements in computational hardware and software allowed for the use of density functional theory based potential energy surfaces (PES) and high performance computing based molecular dynamics simulations. The physics and mathematics behind these approaches is well beyond the scope of this thesis, but the basic ideas behind them could prove useful.

2.1 Density Functional Theory

Density Functional Theory (DFT) is an approach often used to study molecular interactions in many-body systems. Being a computation physics based approach, DFT is a popular way to estimate the interaction of atoms and electrons in a lattice

as functionals [5], or functions of functions. When applied to cases like calculation of many body potentials, the DFT approach eliminates the need for higher order parameters including material properties. This results in what is known as an *ab initio* or first principles calculations [14].

Popular DFT techniques provide interatomic potentials as a combination of those determined by the structure and composition of the unit cell and those potentials that are a result of interatomic interactions. This results in a supercell being represented as a set of n one-electron Schrödinger-like equations, n being the number of atoms in the super cell [14]

2.2 Ab Initio Calculations

DFT techniques estimate the chemical and physical properties of a unit cell or supercell by trying to create functionals that describe the interaction between atoms of that cell. The accuracy of this technique suffers due to issues that are difficult and are often impossible to resolve. These may include difficult-to-estimate functionals that correspond to self interaction of atoms as well as interaction between atoms that are not in the same vicinity. Factors like poor estimation of ionization energy and lack of accurate descriptions of charge transfer within the cell add to the inaccuracy as well [4].

Ab Intio, or first principles calculations seeks to get around this problem by describing the system of n atoms in terms of n one-electron Schrödinger-like equations, or Kohn-Sham equations. Essentially, the properties of the system are estimated by considering the interactions of any atom with its closest neighbours. This approach has reasonably high accuracy provided the system scale, i.e the number of atoms is not too large. The reduction in computational requirements even with today's tech-

nology is an added bonus.

2.3 Supercell Structure Description

For the scope of this thesis, the diamond-like structure from [11] was used. The lattice parameters were according to the table below.

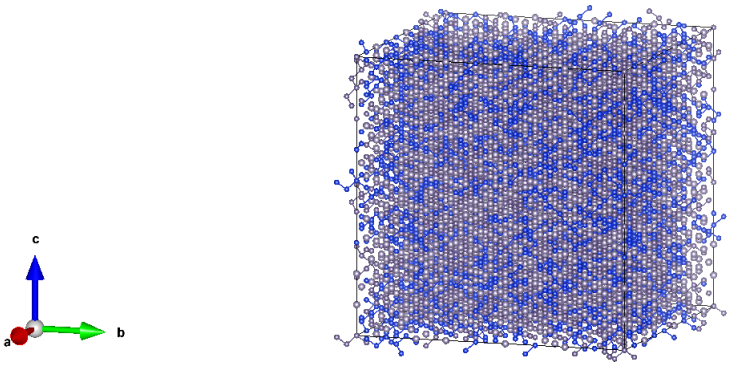

Figure 2.1: Supercell Structure - Vesta

Parameter	**Value**
a	48.87630
b	48.87630
c	48.87630
α	90.0000
β	90.0000
γ	90.0000

Table 2.1: Cell Parameters

The cell contains a total of 5832 atoms, with a volume of 116760.242173 $Å^3$. For the scope of the thesis, the supercell contains three different atoms - silicon, germanium and tin. The concentration of tin was varied as a fraction of the total number of atoms in the supercell. Once the fraction of tin was selected, the remaining atoms were split equally between silicon and germanium.

Chapter 3

MACHINE LEARNING BASED MODELS

Accuracy of results from conventional molecular dynamics (MD) and Monte Carlo (MC) simulations depend on the accuracy of the potential energy surface (PES) needed. The most common ways of calculating PESs are Ab initio methods based on density functional theory (DFT). All this accuracy however, comes with a price. The simulations from these PESs can be carried for a time-frame of a few tens of picoseconds and for a few thousand atoms at best [2]. These simulations are also very expensive in terms of computational power needed. The empirical potentials obtained through these methods are very unreliable when the system is scaled up, in terms of the time-frame and the number of atoms.

3.1 Potential Solutions

One of the most promising ways to tackle this accuracy versus scale problem is by the use of supervised machine learning concepts like neural networks. Running machine learning based algorithms have become more plausible and computationally viable over the past few years. This can be attributed in no small part to development in dedicated hardware and software environments specifically designed for this purpose. As a result, there has been great progress in the field of using machine learning algorithms to calculate DFT based PESs [2; 6; 8; 12]

The idea is to build a network that consumes a certain input; in this case the generalized co-ordinates that determine the energy configuration as the output. The input and output "layers" are associated with each other by means of hidden layers.

The number of hidden layers and the number of nodes per layer determine the accuracy as well as the computational cost of the model. These associations are calculated in terms of weighted real value parameters and activation functions, initially chosen at random. By using a data set comprising of the input and output, the model is "trained" by iterative adjustment of these parameters. The model is then fed fresh data, which then outputs a set of energy configurations, thus resulting in a reasonably accurate PES.

This approach all but removes the need to directly generate a DFT based PES. The hidden layers, by definition is a black box. There is no need for the workings of these layers to be explored. Here, the model "predicts" the output of a DFT based calculation without actually performing them. This has the tremendous advantage of drastically reducing computational requirements as well as allowing for the scaling up of the system both in terms of time scale and the number of atoms. The accuracy of such a system depends on the activation functions used as well as the accuracy of the data set used to "train" the model to begin with.

3.2 DeepMD-Kit

One of the most well known and robust software packages that can be used to train and test a deep learning based NN model is DeepMD-kit [13]. This is a package written in Python/C++ that provides deep learning based PESs to run MD and MC simulations. This approach provides PESs for small scale and high scale systems. And is not limited to pure substances either. This package makes use of TensorFlow, resulting in a very streamlined, efficient and automated training process. DeepMD-kit also integrates well with conventional MD simulations systems like LAMMPS. This results in a system that is a full end-to-end solution for MD simulations, starting from

training of data to get PESs and then running simulations based on the resulting PESs.

The DeepMD method is integrated with LAMMPS using the "pair-style" command in the LAMMPS input file. A trained model was created using a json file containing the input parameters for silicon. The training was conducted with 2000000 iterations, in batches of 100000. The cut-off radius was set at 6Å. After 2000000 iterations were complete, the model was frozen and tested. This model was then passed to LAMMPS as an input using the "pair-style" command.

Chapter 4

SIMULATIONS AND RESULTS

4.1 Simulations in LAMMPS

The super cell discussed in [11] was then built in Vesta and a corresponding POSCAR file was generated. This file contained a description of the cell. More specifically, the number, types and locations of various atoms inside the unit cell. A custom MATLAB script was used to then generate different POSCAR files, one each for different concentrations of tin. The concentration of tin was varied as a fraction of the total number of atoms in the unit cell, ranging from 1/3 to 1/108. These POSCAR files were then converted into a format that can be used as inputs for LAMMPS.

A custom python script was written in order to permute the POSCAR files. The idea here was to randomly shuffle the positions of the atoms within the unit cell in order to account for the limitations of current manufacturing methods. Ten such POSCAR files were created, there by resulting in 10 different configurations or arrangements of atoms in the super cell. These configurations were then used as input data files for LAMMPS and then simulations were carried out to estimate the thermal conductivity of the super cell. The average value was considered as the final thermal conductivity at that particular temperature. These simulations were then repeated for different temperatures, starting from 300K and going up to 1200K with 50K increments. The process of running ten simulations per value of temperature between 300K and 1200K was then repeated for various concentrations of tin. This allowed for the following two separate estimations -

1. Thermal conductivity v/s temperature at constant Sn concentration

2. Thermal conductivity v/s Sn concentration at constant temperature

4.2 Results

For each concentration of tin, a temperature range of 300K – 1200K was selected. There were 10 configurations / arrangements of atoms in the super cell. This resulted in 190 simulations run for each concentration of tin. This provided adequate number of data points to understand the variation of thermal conductivity.

4.2.1 Variation With Atomic Arrangement and Temperature

Tin Concentration - 1/108 Fraction

There was no general trend in the variation of thermal conductivity with respect to atomic arrangement. The value did vary, but in an unpredictable way. For instance, at a temperature of 300K, the following variation was observed.

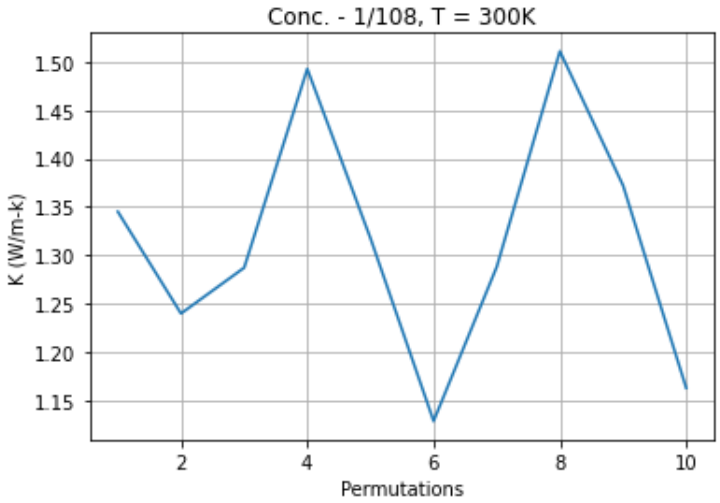

Figure 4.1: Thermal Conductivity v/s Configurations - 300K

The thermal conductivity varied by roughly 15% from the average at this temperature. The unpredictable variation of thermal conductivity was demonstrated by the

lack of a general trend in the plots below.

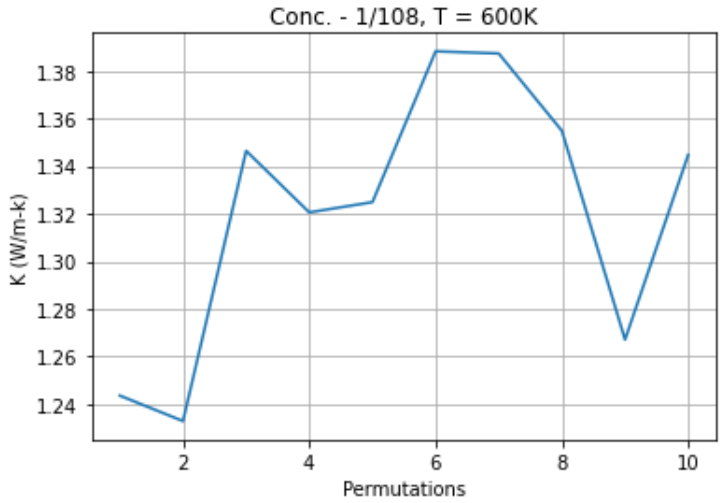

Figure 4.2: Thermal Conductivity v/s Configurations - 600K

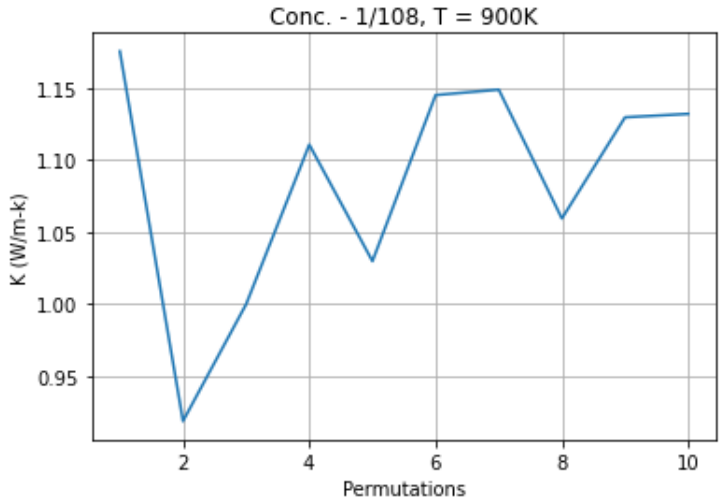

Figure 4.3: Thermal Conductivity v/s Configurations - 900K

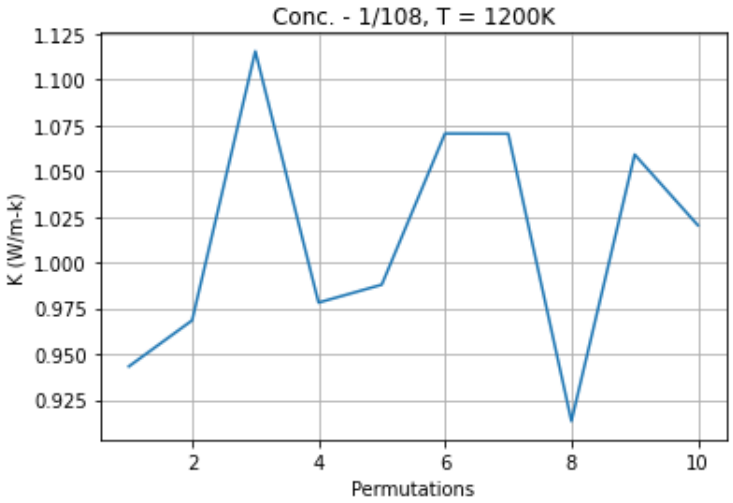

Figure 4.4: Thermal Conductivity v/s Configurations - 1200K

Once all 190 simulations were run, the average thermal conductivity was calculated at each temperature. A plot was then drawn comparing the average temperature against temperature and the following was obtained.

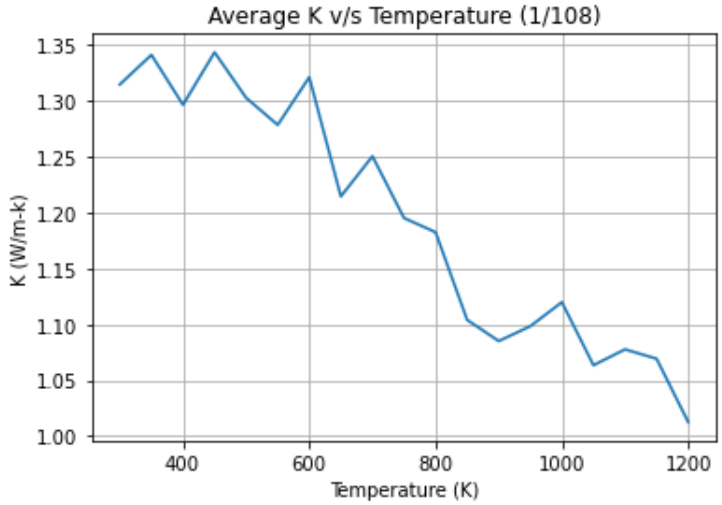

Figure 4.5: Average Thermal Conductivity v/s Temperature

It was observed from this plot that the average thermal conductivity had an overall downward trend with increase in temperature. Thermal conductivity in metals and alloys is an electronic effect and depends very strongly on the flow of free electrons. At higher temperatures, the atoms in the super cell vibrate with greater frequency

and amplitude. This obstructs the smooth flow of free electrons, thereby decreasing the overall thermal conductivity. This effect was seen in the super cell used in the simulation as well in the form of the overall downward trend.

The variation of thermal conductivity from the average can be visualized using the error bar below.

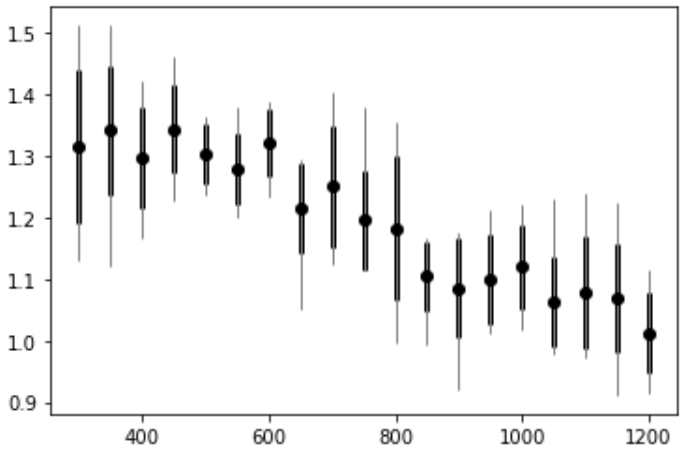

Figure 4.6: Average Thermal Conductivity - Error Bar

The table below shows the percentage variation of the maximum and minimum values of thermal conductivity from the average value.

Temperature	**Variation of max. K (%)**	**Variation of min. K (%)**
300K	14.949119947731823	14.149236310133576
350K	12.7214350908812	16.51837827154006
400K	9.544779573975525	10.138664747497927
450K	8.628067878692104	8.573220246066182
500K	4.726078451850939	5.163781056781602
550K	7.82324763624621	6.2086976650497325
600K	5.0874089208085715	6.676440621681132
650K	6.433412664163807	13.450336440918974
700K	12.169224657549503	10.031526136749063

750K	15.42051472254751	6.7305970330663865
800K	14.662706297342986	15.680670566455554
850K	5.743987799151685	9.983079764446083
900K	8.339562683542974	15.313812712768806
950K	10.244581471337241	7.945156773677783
1000K	8.976832154987289	9.281153405119532
1050K	15.563767334433413	8.195659603475299
1100K	15.056197084917205	9.834027072332102
1150K	14.46986554505234	14.779647811091593
1200K	10.141777483415229	9.800756466442328

Table 4.1: Variation of K From Average Value (Sn - 1/108)

From the table, it was observed that the highest variation occurred at 800K while the lowest variation occurred at 500K.

Tin Concentration - 1/12 Fraction

The POSCAR file were rebuilt by increasing the concentration of tin to 1/12 fraction. This file was then put through 10 different rearrangements to create 10 configurations and the simulations were re-run. Here again, no general trend in the variation of thermal conductivity was observed at any particular temperature. The following plots show this lack of trend at 300K, 600K, 900K and 1200K.

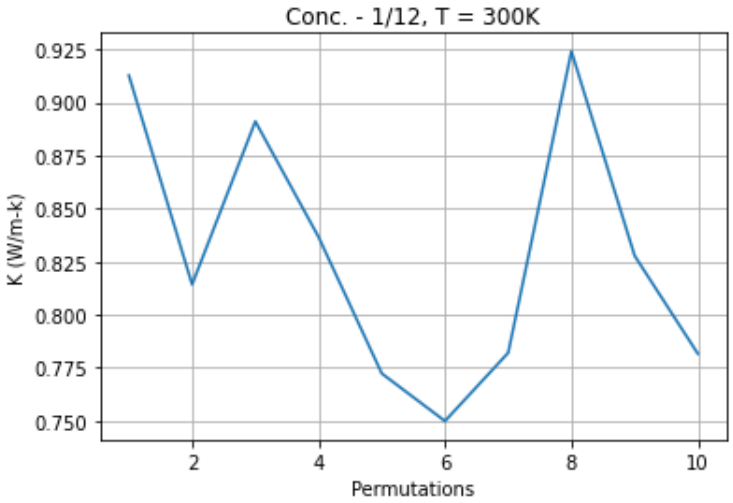

Figure 4.7: Thermal Conductivity v/s Configurations - 300K

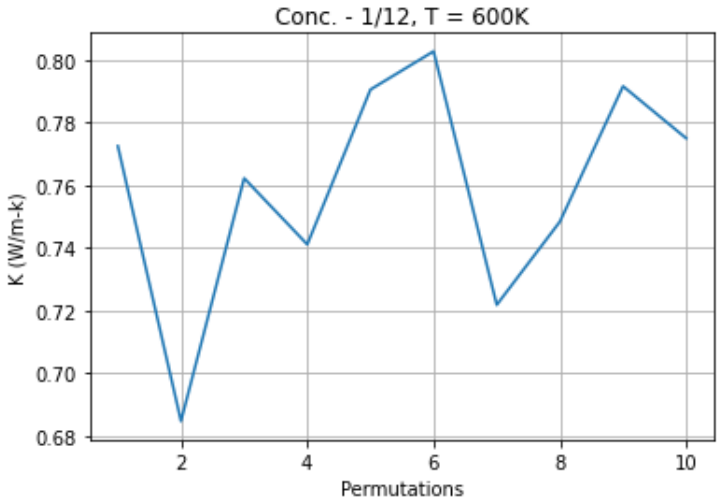

Figure 4.8: Thermal Conductivity v/s Configurations - 600K

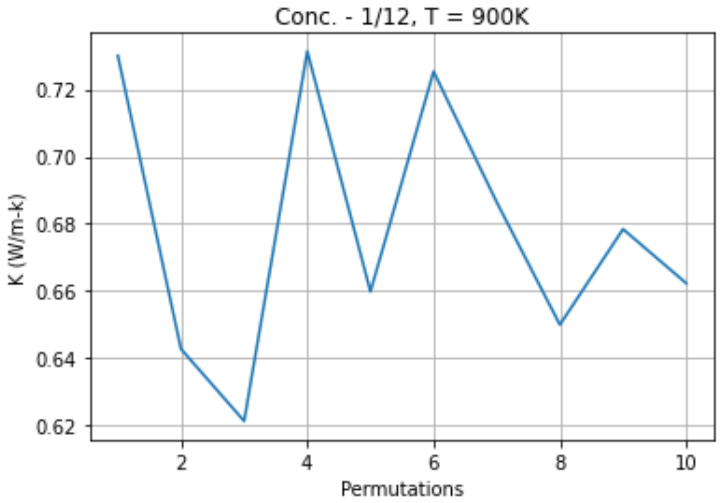

Figure 4.9: Thermal Conductivity v/s Configurations - 900K

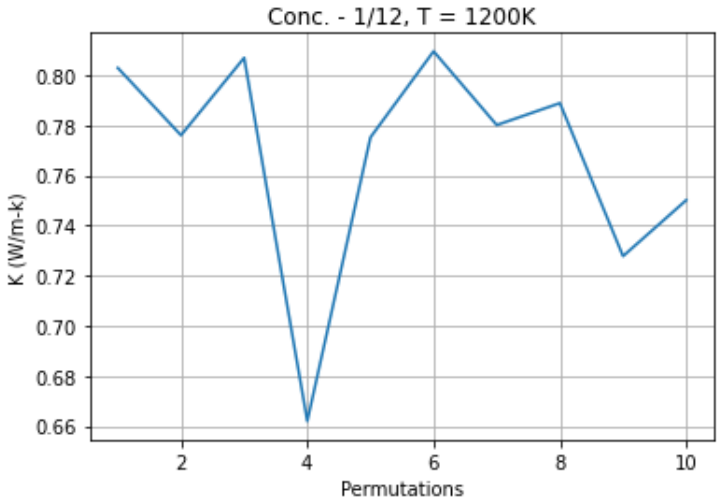

Figure 4.10: Thermal Conductivity v/s Configurations - 1200K

There was a difference in the trend of the average thermal conductivity against temperature compared to the 1/108 fraction. While in the case of the latter there was a clear downward trend, in this case there was no trend at all. This can be seen in the plot below.

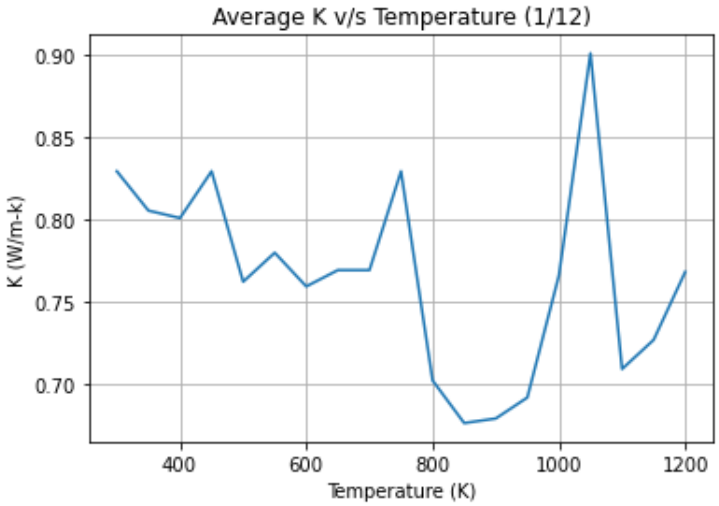

Figure 4.11: Average Thermal Conductivity v/s Temperature

The following error bar was plotted, showing from the variation of the average thermal conductivity at each particular temperature.

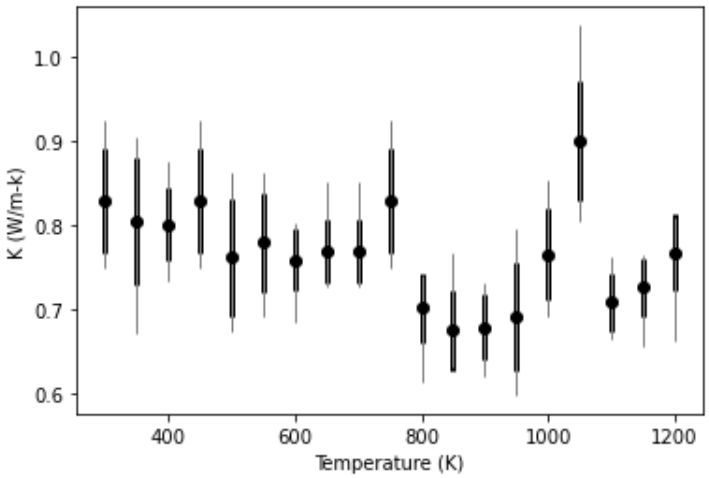

Figure 4.12: Average Thermal Conductivity - Error Bar

The table below shows the maximum and minimum variations of thermal conductivity from the average at each particular temperature.

Temperature	**Variation of max. K (%)**	**Variation of min. K (%)**
300K	11.4274134612636	9.55889495891503
350K	12.337150116168978	16.62134551525248

400K	9.499692398778114	8.408502248956541
450K	11.4274134612636	9.55889495891503
500K	13.355471887180483	11.623731904065792
550K	10.777028441441377	11.275333418496803
600K	5.752647104200773	9.788890690298286
650K	10.693704566486732	5.525875934271937
700K	10.693704566486732	5.525875934271937
750K	11.4274134612636	9.55889495891503
800K	5.634530644956691	12.45707591165573
850K	13.482039254282341	6.593814005433288
900K	7.759071955954075	8.483419001574678
950K	15.086183426397549	13.45393964642916
1000K	11.42456236066247	9.566524843597072
1050K	15.178030072110577	10.543298767345743
1100K	7.671657975845833	6.229199147354027
1150K	5.283294044696646	9.623410450628976
1200K	5.414211346209397	13.793854251095174

Table 4.2: Variation of K From Average Value (Sn - 1/12)

From the table, it was determined that the highest variation occurred at 350K while the lowest occurred at 1100K.

Tin Concentration - 1/3 Fraction

Similar to other tin concentrations, there was no general trend in the variation of thermal conductivity with respect to the atomic configuration at any particular tem-

perature. The corresponding trends at 300K, 600K, 900K and 1200K were seen.

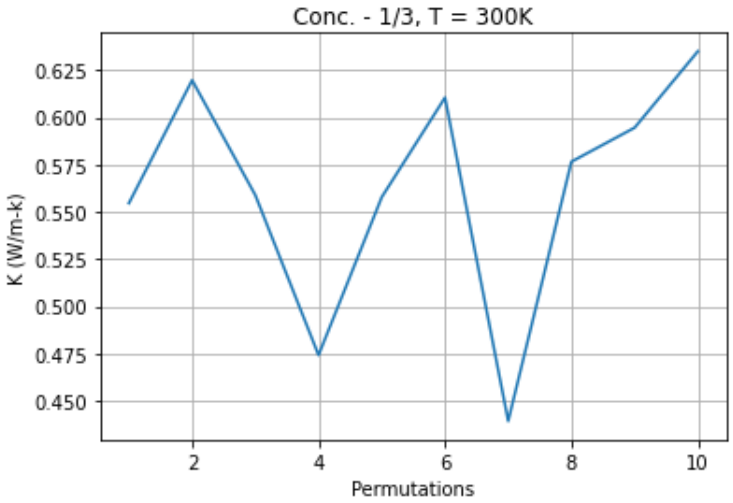

Figure 4.13: Thermal Conductivity v/s Configurations - 300K

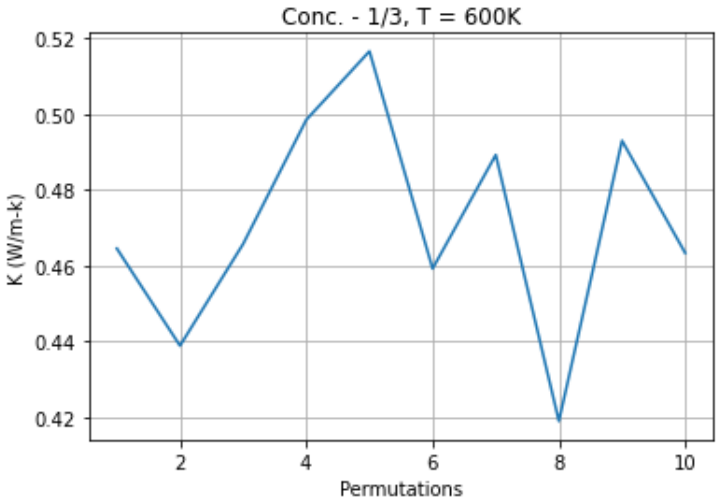

Figure 4.14: Thermal Conductivity v/s Configurations - 600K

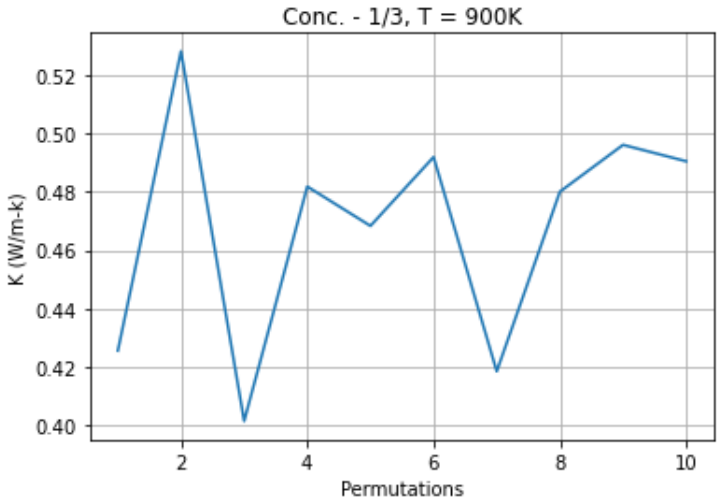

Figure 4.15: Thermal Conductivity v/s Configurations - 900K

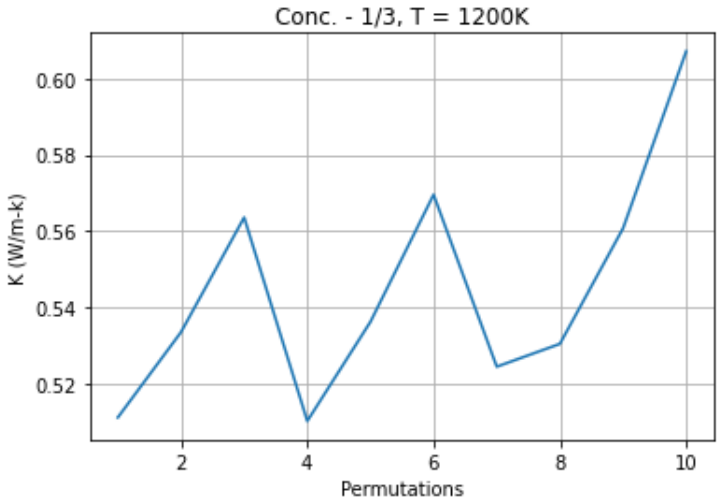

Figure 4.16: Thermal Conductivity v/s Configurations - 1200K

However, there was a marked difference observed in the trend in the variation of the average thermal conductivity with temperature. While at 1/108 fraction the thermal conductivity decreased overall with temperature, in this case, the thermal conductivity first decreased, then increased again after reaching a minimum value at 700K. This trend can be observed from the following plot.

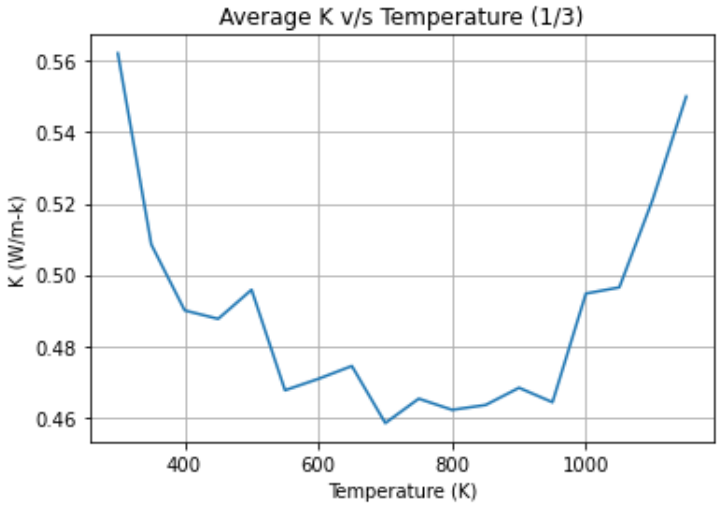

Figure 4.17: Average Thermal Conductivity v/s Temperature

The following error bar shows the variation from the average value. Table 4.2 shows the percentage deviation of K from the average value at each temperature.

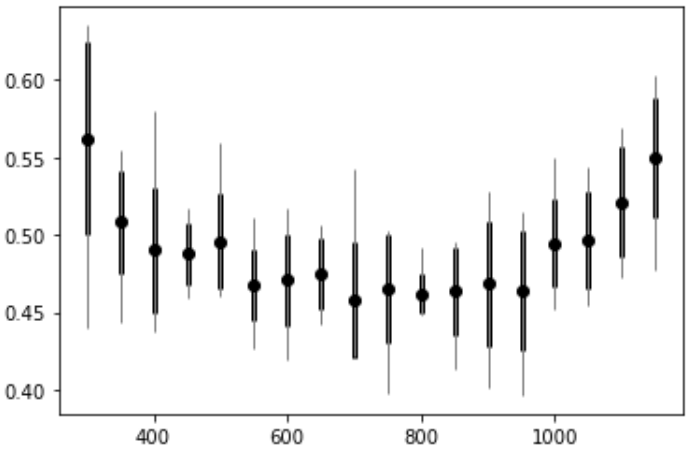

Figure 4.18: Average Thermal Conductivity - Error Bar

Temperature	**Variation of max. K (%)**	**Variation of min. K (%)**
300K	12.958013488319297	21.83436039486742
350K	9.001024907207986	12.872964319651881
400K	18.239729085816073	10.714939408329462
450K	6.014510451968916	5.754531464472458

500K	12.689905252466255	7.103916420967426
550K	9.302644919922038	8.836434616728816
600K	9.736595613902518	11.020126538493201
650K	6.803550977199705	6.859851413046127
700K	18.36498281695152	8.027514905231497
750K	8.052622715203176	14.410490904833548
800K	6.285332825926446	3.065405248764371
850K	7.000122887413854	10.689567308840457
900K	12.774838499859667	14.238824034108067
950K	10.868949283971356	14.582026001512657
1000K	11.190456126456686	8.773132565615798
1050K	9.478799200566733	8.437931427054918
1100K	9.092398202564251	9.443881600873315
1150K	9.563884047164375	13.145725960008765

Table 4.3: Variation of K From Average Value (Sn - 1/3)

The highest deviation was found to occur at 300K, while the lowest was observed at 800K.

Chapter 5

FUTURE WORK

Given the variations of thermal conductivity, a crucial next step appears to be to run conventional simulations in LAMMPS for a wider dataset, including different concentrations of tin. It is proposed that these simulations are repeated for tin concentrations of 1/4, 1/6, 1/9, 1/18, 1/27, 1/36 and 1/54.

It is also proposed that LAMMPS simulations are run for pure silicon and germanium using the trained model for DeepMD-Kit. The final step would be to recreate models for the diamond structure with varying tin concentrations, train individual models for these structures in DeepMD-kit and then run MD simulations in LAMMPS to see how closely the results match.

REFERENCES

[1] *Annales de chimie et de physique*, no. v. 56 in Annales de chimie et de physique (1834), URL https://books.google.com/books?id=1Jc5AAAAcAAJ.

[2] Behler, J. and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces", Phys. Rev. Lett. **98**, 146401, URL https://link.aps.org/doi/10.1103/PhysRevLett.98.146401 (2007).

[3] Fugallo, G. and L. Colombo, "Calculating lattice thermal conductivity: a synopsis", Physica Scripta **93** (2018).

[4] Grabowski, I., A. Teale, E. Fabiano, S. Śmiga, A. Buksztel and F. Della Sala, "A density difference based analysis of orbital–dependent exchange–correlation functionals", Molecular Physics **112**, 700 (2014).

[5] Grant, I., "Density functional theory", URL http://www.chm.bris.ac.uk/webprojects2002/grant/webcomp/dft.html (2002).

[6] Huan, T. D., R. Batra, J. Chapman, S. Krishnan, L. Chen and R. Ramprasad, "A universal strategy for the creation of machine learning-based atomistic force fields", npj Computational Materials **3**, 1, 37, URL https://doi.org/10.1038/s41524-017-0042-y (2017).

[7] Kok S Ong, K. C. L., Liben Jiang, *Thermoelectric Energy Conversion* (Elsevier Inc., Amsterdam, Netherlands, 2018).

[8] Rupp, M., A. Tkatchenko, K.-R. Müller and O. A. von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning", Phys. Rev. Lett. **108**, 058301, URL https://link.aps.org/doi/10.1103/PhysRevLett.108.058301 (2012).

[9] Sadagdar, M., "Making a particle detector (cloud chamber)", URL https://youtu.be/eh3bxXHqF2U (2019).

[10] Tritt, T., *Thermal Conductivity: Theory, Properties, and Applications*, Physics of Solids and Liquids (Springer US, 2005), URL https://books.google.com/books?id=whJNfKmziiIC.

[11] Wang, D., L. Liu, M. Chen and H. Zhuang, "Electrical and thermal transport properties of medium-entropy siygeysnx alloys", Acta Materialia **199**, 443–452, URL https://www.sciencedirect.com/science/article/pii/S1359645420306601 (2020).

[12] Wang, H., L. Zhang, J. Han and W. E, "Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics", Computer Physics Communications **228**, 178–184 (2018).

[13] Wang, H., L. Zhang, J. Han and W. E, "Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics", Computer Physics Communications **228**, 178–184, URL https://www.sciencedirect.com/science/article/pii/S0010465518300882 (2018).

[14] Wikipedia contributors, "Density functional theory — Wikipedia, the free encyclopedia", URL https://en.wikipedia.org/w/index.php?title=Density$_f$unctional$_t$heoryoldid = 1010216859, [$Online; accessed 3 - April - 2021$](2021).

APPENDIX A

RAW DATA

DATA COLLECTED FROM LAMMPS SIMULATIONS, BETWEEN DECEMBER 2020 - APRIL 2021

Tin Concentration = 1/3

Results_1_3

Temperature (K) / Configuration	300	350	400	450	500
	Thermal Conductivity (W/m-K)				
1	0.5546159901	0.443036870255	0.472651788441	0.473144715634	0.460636824298
2	0.619678527526	0.520629977163	0.531557926777	0.482584541803	0.50947710773
3	0.558872255922	0.533671369434	0.4970183085	0.480600007001	0.495607172728
4	0.474261298334	0.503234101982	0.496006784481	0.50632843385	0.493877528611
5	0.557828695785	0.505768512752	0.478808627235	0.464930028736	0.493014313381
6	0.610450339628	0.554265074582	0.468338903833	0.459550959559	0.466155103817
7	0.439383157562	0.48722957906	0.437477636885	0.490699550051	0.491295611131
8	0.576568616063	0.480372127508	0.479412578709	0.512982930451	0.558693090582
9	0.594564105196	0.54747089723	0.579349299016	0.488347458606	0.460559443381
10	0.634957366039	0.509274361384	0.459163552393	0.516938063574	0.528475625448

Temperature (K) / Configuration	550	600	650	700	750
	Thermal Conductivity (W/m-K)				
1	0.460275988823	0.464507545204	0.474710957923	0.477944995963	0.478376370623
2	0.475577834151	0.438818482529	0.449492895462	0.441024714785	0.480046263946
3	0.426264440187	0.465729686897	0.46479187851	0.467868979686	0.479390045855
4	0.511079514627	0.49852777965	0.505477982691	0.427413004444	0.461502560279
5	0.459288268589	0.516630687325	0.451905190033	0.428811080222	0.502680808455
6	0.457416892876	0.459202650293	0.490363842309	0.542563100893	0.502158904331
7	0.458612049763	0.48924742378	0.441844475772	0.467682848478	0.485746093043
8	0.489806538425	0.418909780528	0.473251968997	0.421584792413	0.420196475922
9	0.475748439121	0.493006735953	0.506661839261	0.428335546842	0.398178244508
10	0.461750229751	0.463335346751	0.485365953101	0.48058523659	0.443909398223

Temperature (K) / Configuration	800	850	900	950	1000
	Thermal Conductivity (W/m-K)				
1	0.460617178568	0.426898944515	0.425688724356	0.401872384298	0.495560514463
2	0.457015068755	0.47968563774	0.528100072981	0.396528935676	0.5056642027
3	0.449917340118	0.49125948896	0.401600958946	0.464957823427	0.511894878433
4	0.455648370709	0.495884844815	0.481832231413	0.514677934879	0.505190478301
5	0.458003337815	0.477529610482	0.468372066159	0.474432776623	0.550074988868
6	0.467120547157	0.413903170018	0.491929646924	0.478956955424	0.461967200215
7	0.474159555942	0.463511417424	0.418614454731	0.460643328216	0.468441688994
8	0.491133153184	0.485796300894	0.480071709009	0.486575229228	0.50866753211
9	0.44792439283	0.455981930612	0.496093223702	0.460656785483	0.488368357742
10	0.45935408057	0.443981133994	0.490479702744	0.502916774615	0.451312278379

Temperature (K) / Configuration	1050	1100	1150	1200
	Thermal Conductivity (W/m-K)			
1	0.459081040764	0.476016018613	0.601070413293	0.511182110925
2	0.543499637665	0.568518125475	0.602596426539	0.533531562439
3	0.454553314949	0.530852130855	0.514418668594	0.563583690382
4	0.482057745002	0.543523995939	0.534197003854	0.510238651527

5	0.53343361399	0.51839918398	0.571806610988	0.536250451886
6	0.465857165681	0.494354451741	0.545094696214	0.569555276337
7	0.495637357911	0.471919176137	0.528913119	0.524483038404
8	0.494386743562	0.558830448444	0.570685691162	0.530462781117
9	0.518197776875	0.493477316674	0.553477727026	0.56074186432
10	0.517723800053	0.555454166479	0.47769459454	0.607114398947

Tin Concentration = 1/12

Results_1_12

Temperature (K) / Configuration	300	350	400	450	500
	Thermal Conductivity (W/m-K)				
1	0.912635472122	0.904645993306	0.810829362235	0.912635472122	0.777363577158
2	0.814232196924	0.799333472089	0.876822333082	0.814232196924	0.848482185019
3	0.891011700978	0.808016275374	0.799373737794	0.891011700978	0.779882302021
4	0.836624699262	0.883627521392	0.824231522978	0.836624699262	0.692872118218
5	0.772369230352	0.765172517004	0.782383322862	0.772369230352	0.709780038167
6	0.749974261887	0.671444536637	0.753969698646	0.749974261887	0.684725922069
7	0.782105331881	0.721603796657	0.836080848025	0.782105331881	0.767493257015
8	0.924001228497	0.825388598001	0.761675430746	0.924001228497	0.822214740904
9	0.827780305201	0.896221458903	0.828744244625	0.827780305201	0.673428269907
10	0.781670464967	0.77750064007	0.733421884475	0.781670464967	0.86377011569

Temperature (K) / Configuration	550	600	650	700	750
	Thermal Conductivity (W/m-K)				
1	0.777731372027	0.772552862942	0.775241654738	0.775241654738	0.912635472122
2	0.860589767432	0.684825956129	0.761943185555	0.761943185555	0.814232196924
3	0.70881306617	0.762307277485	0.794199491693	0.794199491693	0.891011700978
4	0.863651992036	0.741190290514	0.765232315059	0.765232315059	0.836624699262
5	0.821920500301	0.790606704204	0.778360916083	0.778360916083	0.772369230352
6	0.6917249552	0.802807528036	0.851298833143	0.851298833143	0.749974261887
7	0.724352633901	0.721907078961	0.778066791717	0.778066791717	0.782105331881
8	0.782489853968	0.748457648871	0.732316300604	0.732316300604	0.924001228497
9	0.774377315309	0.791660875537	0.726560845482	0.726560845482	0.827780305201
10	0.790657978807	0.775054301077	0.727360051477	0.727360051477	0.781670464967

Temperature (K) / Configuration	800	850	900	950	1000
	Thermal Conductivity (W/m-K)				
1	0.655704288581	0.713513970856	0.730076170367	0.776511422884	0.709498352648
2	0.614443621188	0.729013357429	0.642627139719	0.717582871342	0.853232746248
3	0.713853537937	0.76709184866	0.621136116693	0.795817658178	0.692493654233
4	0.679887989613	0.701066253289	0.731376224537	0.615164463481	0.79970636341
5	0.696968028637	0.631387348709	0.659741445736	0.647094145001	0.84105023855
6	0.717714636545	0.639924837	0.72535887424	0.710672889147	0.748204094806
7	0.741424440729	0.648938436111	0.686378468519	0.660451664536	0.797257967865
8	0.740814343305	0.649669691571	0.649860388144	0.598463525547	0.722667709305
9	0.721370423071	0.636704853334	0.678355615479	0.702719764443	0.758901355412
10	0.736588369231	0.642277560786	0.662232497403	0.690492919405	0.734479978532

Temperature (K) / Configuration	1050	1100	1150	1200
	Thermal Conductivity (W/m-K)			
1	1.03781290148	0.711639245896	0.728762159958	0.802864550406
2	0.923737783129	0.664698001043	0.710333485704	0.776036612237
3	0.823153936809	0.671252089551	0.756789098656	0.806914105844
4	0.948707173477	0.735019007836	0.764975747476	0.662005944788

5	0.942703643455	0.746498326188	0.741886186345	0.775174185212
6	0.905875431558	0.693983860634	0.764393995179	0.809511131371
7	0.882446242692	0.666031574195	0.703735943507	0.780087343393
8	0.823570332679	0.709753187591	0.736169509265	0.788835891434
9	0.91645292121	0.726429170974	0.656665426098	0.727793204312
10	0.806050586253	0.763234772176	0.702168133826	0.750112872281

Tin Concentration = 1/108

Results_1_108

Temperature (K) / Configuration	300	350	400	450	500
	Thermal Conductivity (W/m-K)				
1	1.34517084376	1.4103377345	1.23443419514	1.41769534179	1.23532714256
2	1.23982147639	1.30407079358	1.21921215499	1.28841191386	1.32724217883
3	1.28720470195	1.36372925169	1.30540573602	1.35935416038	1.31983197279
4	1.49286668402	1.3629718792	1.42012136637	1.37688055703	1.35469616618
5	1.31774915481	1.43408628582	1.37129490381	1.31131966214	1.24747625863
6	1.12857039636	1.30527473182	1.29543840807	1.30925505175	1.27008737661
7	1.28810580534	1.31407277825	1.40228535715	1.40079681584	1.25702188345
8	1.51108934021	1.51159149178	1.25493853658	1.22813476715	1.35332889603
9	1.37232490997	1.28435249737	1.29576445302	1.28193953322	1.36415146752
10	1.16282012463	1.11948636054	1.16494827685	1.45919945129	1.29673706049

Temperature (K) / Configuration	550	600	650	700	750
	Thermal Conductivity (W/m-K)				
1	1.30458699684	1.24355674305	1.17373995568	1.2605595424	1.18372005446
2	1.23857147372	1.23286888219	1.18509232036	1.25654821832	1.3794053649
3	1.30717618447	1.3464468588	1.26380909948	1.40249609784	1.22850536322
4	1.37838995082	1.32056696838	1.22591729335	1.12491134629	1.19399949564
5	1.35120268989	1.32488343166	1.05107990428	1.31992288288	1.26562203991
6	1.1990084833	1.38827748568	1.26553944778	1.14066622254	1.13316126496
7	1.23393723864	1.38736244585	1.29255292967	1.15007946606	1.1146745892
8	1.22498983091	1.35486880396	1.26882767167	1.23234557608	1.16929352043
9	1.2614222495	1.26711561287	1.2600777942	1.22394264978	1.11842151135
10	1.28450671896	1.34474565586	1.15760379265	1.39192277585	1.16432496596

Temperature (K) / Configuration	800	850	900	950	1000
	Thermal Conductivity (W/m-K)				
1	1.26846035624	1.10311929433	1.17555037466	1.10635745338	1.12456258368
2	1.35575808448	1.11566165814	0.918896815975	1.05177170093	1.16169191844
3	1.19913401535	1.10451504042	1.00001491705	1.01109288775	1.07307472387
4	1.01906156721	1.02885842525	1.11075640794	1.21088156073	1.22035833653
5	1.07321731877	1.16239516425	1.02989222593	1.1196635922	1.01589942131
6	1.19298308477	0.993769453322	1.14516504646	1.19903686416	1.11520827305
7	1.24301852115	1.16739324865	1.14895554389	1.07911142115	1.04728428045
8	0.996981636392	1.15634722137	1.0595122645	1.15620638065	1.18465363696
9	1.19009692961	1.0848713258	1.12975467691	1.0125022156	1.06956707817
10	1.28516852741	1.12287647049	1.13211202584	1.03696844627	1.18602798298

Temperature (K) / Configuration	1050	1100	1150	1200
	Thermal Conductivity (W/m-K)			
1	1.03559132166	1.10794204061	1.08499757442	0.943250264884
2	1.06783791317	1.03984398226	1.061303138	0.968377577417
3	1.10814162162	1.23989467562	0.911198427682	1.11533844156
4	0.976299146104	0.97166699915	1.00340852186	0.978094358399

5	1.07133424317	1.02946028427	1.03246365145	0.987876066453
6	1.10117405207	1.09163466436	1.22394192024	1.07042301187
7	1.04816612643	1.21102487869	1.09882498661	1.07035088388
8	1.22897029576	1.09440523595	1.0186465693	0.913392592813
9	1.01825012006	0.977027741515	1.07764640179	1.05894327214
10	0.978799298069	1.01352620934	1.17983191573	1.02034214573