UNIVERSITY OF TURKU

Investigating atomic mobility in phase-change memory materials with machine-learned molecular-dynamics simulations

Materials Engineering/Faculty of Technology
Master's thesis
Department of Mechanical and Materials Engineering

Author:
Ville Paakkola

Supervisors:
Academy Research Fellow Konstantinos Konstantinou
Associate Professor Milica Todorović

22.12.2025
Turku

Master's thesis

Subject: Materials Engineering
Author: Ville Paakkola
Title: Investigating atomic mobility in phase-change memory materials with machine-learned molecular-dynamics simulations
Supervisors: Academy Research Fellow Konstantinos Konstantinou and Associate Professor Milica Todorović
Number of pages: 65 pages
Date: 22.12.2025

Phase-change materials are materials that can change between at least two structurally different solid phases, an amorphous (glass) and crystalline phase. A functionally important aspect is that the properties of these two phases differ from each other. This is due to the differences in the atomic structures of these two phases. The current and possible future applications of phase change materials are based on the properties of the two structural states.

In this thesis, the short-range order structure and ionic diffusion properties were studied for a ternary germanium-antimony-tellurium phase-change memory material model, in the composition of $Ge_2Sb_2Te_5$, in three different temperature regimes by using molecular-dynamics simulations with a machine-learned interatomic potential. Studied properties included radial distribution functions, coordination numbers, bond angle distributions, mean squared displacements and diffusion coefficients. The study was conducted by first defining the atomistic model for the simulations, and then defining the simulation environment and settings. After that, the simulations were performed by using high performance computing facilities, and lastly the simulation results were obtained and analysed.

The results of the thesis provided information about the local atomic structure and ionic diffusion in the amorphous $Ge_2Sb_2Te_5$ phase-change memory material, in three distinct temperatures. The radial distribution functions provided information about the arrangement of the atoms in the structure, the coordination numbers about the local atomic environments, and the bond angle distributions about the local geometry of the first coordination shell. The mean squared displacements provided insights into the mobility of the different atoms in the structure, while the diffusion coefficients indicated the rate of transfer of particles in given unit area.

Key words: Phase-change memory materials, Molecular-dynamics simulations, Machine learning, Mean squared displacement, Glasses, High performance computing

Diplomityö

Oppiaine: Materiaalitekniikka
Tekijä: Ville Paakkola
Otsikko: Atomien liikkuvuuden tutkiminen faasimuutosmuistimateriaaleissa koneoppineen molekyylidynamiikan simulaation avulla
Ohjaajat: Akatemiatutkija Konstantinos Konstantinou ja apulaisprofessori Milica Todorović
Sivumäärä: 65 sivua
Päivämäärä: 22.12.2025

Faasimuutosmateriaalit ovat materiaaleja, jotka voivat vaihtaa ainakin kahden rakenteellisesti erilaisen kiinteän faasin, amorfisen (lasi) ja kiteisen faasin välillä. Toiminnallisesti tärkeä näkökohta on se, että näiden kahden faasin ominaisuudet eroavat toisistaan. Tämä johtuu näiden kahden faasin atomirakenteiden eroista. Faasimuutosmateriaalien nykyiset ja mahdolliset tulevaisuuden sovellukset perustuvat sen kahden rakenteellisen olomuodon ominaisuuksiin.

Tässä diplomityössä tutkittiin lyhyen etäisyyden järjestysrakennetta ja ionidiffuusio-ominaisuuksia ternaarisen germanium-antimoni-telluuri faasimuutosmuistimateriaalimallissa yhdisteessä $Ge_2Sb_2Te_5$, kolmessa eri lämpötila-alueessa käyttämällä molekyylidynamiikan simulaatioita koneoppineen interatomisen potentiaalin kanssa. Tutkittuja ominaisuuksia olivat säteittäiset jakautumafunktiot, koordinaationumerot, sidoskulmajakaumat, keskineliömäiset siirtymät ja diffuusiokertoimet. Tutkimus tehtiin määrittelemällä ensin simulaatioiden atomistinen malli ja sitten simulaatioympäristö ja -asetukset. Tämän jälkeen simulaatiot suoritettiin käyttämällä korkean suorituskyvyn laskentalaitteita, ja lopuksi simulaatiotulokset saatiin ja analysoitiin.

Diplomityön tulokset antoivat tietoa paikallisesta atomirakenteesta ja ionisesta diffuusiosta amorfisessa $Ge_2Sb_2Te_5$ -faasimuutosmuistimateriaalissa kolmessa erilaisessa lämpötilassa. Säteittäiset jakaumafunktiot antoivat tietoa atomien järjestymisestä rakenteessa, koordinaationumerot paikallisista atomiympäristöistä ja sidoskulmajakaumat ensimmäisen koordinaatiokuoren paikallisgeometriasta. Keskineliömäiset siirtymät antoivat tietoa eri atomien liikkumisesta rakenteessa ja diffuusiokertoimet indikoivat hiukkasten siirtymänopeudesta tietyllä yksikköalueella.

Avainsanat: Faasimuutosmuistimateriaalit, Molekyylidynamiikan simulaatio, Koneoppiminen, Keskimääräinen neliösiirtymä, Lasit, Suurteholaskenta

Abbreviations

MD	molecular dynamics
RDF	radial distribution function
CN	coordination number
MSD	mean squared displacement
PCM	phase change material
ML	machine learning
GAP	Gaussian Approximation Potential

Table of contents

1 Introduction 7

1.1 Research questions of the thesis and the motivation 8

1.2 Outline of the thesis 8

2 Phase change materials 10

2.1 Amorphous phase 10

2.2 Crystalline phase 11

2.3 Electronic memory and digital storage 12

2.4 Phase change process 12

2.5 Fragility 13

3 Methods 15

3.1 Computer simulations 15

3.2 Classical molecular dynamics 16

3.2.1 Equations of motion 16

3.2.2 Schemes of integration 18

3.2.3 Interatomic potential 20

3.2.4 Periodic boundary conditions 22

3.2.5 Ensembles 24

3.2.6 LAMMPS 25

3.3 Machine-learning potentials 26

3.3.1 Types of different machine learning interatomic potentials 27

3.3.2 Gaussian Approximation Potential (GAP) 27

3.4 Modelling properties 28

3.4.1 Short-range order structure 28

3.4.2 Ionic diffusion 30

4 Computational details 31

4.1 Input file 31

4.2 Model structure file 36

4.3 CSC High Performance Computing Facilities 37

4.4 Study workflow 38

5 Results 39

5.1 Temperature profiles **39**

5.2 Local atomic structure **41**

5.2.1 Radial Distribution Functions 41

5.2.2 Coordination Numbers 48

5.2.3 Bond Angle Distributions 51

5.3 Atomic mobility **55**

5.3.1 Mean Squared Displacement 55

5.3.2 Diffusion Coefficients 56

6 Discussion and conclusions **60**

7 Outlook **62**

References **63**

1 Introduction

Throughout the history of human development one important aspect has been the capability of storing information. Humans desire to store information and knowledge for the future has been seen in ancient cave paintings, hieroglyphs on pyramids, archival scrolls in the eighth century to medieval books. During the past century, new techniques have seen the development of storing significant amounts of data. In the early days, the first optical storage products were only able to read encoded information. Now since the development has gone further, recordable, and rewritable storage is available. [1]

Phase change materials considered materials which are able to change between at least two solid phases that are structurally distinct from each other, an amorphous and crystalline phase. The important aspect of the functionality of the structural phases is that often the properties of these two phases are very different regarding electrical and optical properties. This is due to the fact that the two phases have differences in their atomic structures. These differences can be then used in different technological applications. [2] The success of rewritable phase change material applications is based on the discovery of metal alloy that contains elements such as germanium (Ge), antimony (Sb) and tellurium (Te) [1]. The information storing is possible due to the switching between the two different phases [2].

Due to the characteristics of phase change materials, they are able to be used in many different applications, for instance, in optical storage products and in electronic memory [2]. In optical storage, phase change materials have been used in compact disks (CD), digital versatile disks (DVD), and Blu-ray disks (BD) during the past three optical storage generations. In electronic memory, phase change materials are showing great potential. The successful usage of phase change materials in current technological applications has fueled their implementation in future applications. [1–3] Lately, companies such as Intel, BAE Systems, Samsung, IBM and STMicroelectronics have started different research & development projects to investigate the phase change random access memory technology [3]. Future applications include in-memory computing, which is a computational approach to tackle processor-memory gaps in computing systems [4]. Also, another possible future application includes phase change material devices for synaptic realization in the field of neuromorphic computing [5].

1.1 Research questions of the thesis and the motivation

To understand at the atomic level the fast crystallization behavior of phase change materials and to use this information to design materials with enhanced crystallization speed, relevant research on this topic is needed. This thesis aims to answer the following research questions:

- How the fast crystallization process is related to the atomic mobility within the structure?
- How different or connected is the dynamical behavior of the atoms among the different temperatures?
- How can dynamical heterogeneity affect the ionic diffusion mechanisms in the atomic structure?

1.2 Outline of the thesis

The thesis is divided into two, logically structured different parts that are literature review and atomistic simulation parts that include a total of seven chapters combined. The literature review part of the thesis consists of chapters two and three where the background and the methods will be discussed. In chapter two, phase change materials and the amorphous and crystalline phases will be described. After that, in the same chapter, applications of phase change materials, electronic memory and digital storage will be discussed. Lastly in chapter two, the crystallization process of these materials will be discussed along with fragility and how it is related to phase change materials.

In the chapter three relevant methods to the thesis will be presented. In chapter three computer simulations will be presented and being more specific, classical molecular dynamics and different aspects that are significantly relevant to it. After that, in the same chapter machine-leaned potentials will be introduced along with the different available types, and the relevant machine-learned potential architecture regarding to this thesis, Gaussian Approximation Potential. As the last thing presented in chapter three, relevant modeling properties that are studied in this thesis will be introduced, in order to first discuss the short- range order structural properties and then move to ionic diffusion properties.

The atomistic simulation part of the thesis includes chapters four and five. In chapter four, computational details regarding simulation will be presented. Firstly, the input file for the

simulations will be introduced and what kind of information the input file included, as well as what the different command lines in the file mean. Then the model structure file will be introduced, what kind of information it holds and what different lines in the model structure file mean. After that CSC high performance computing facilities will be presented, and how the computing facilities can be utilized. Lastly, in chapter four, the study workflow will be discussed, what kind of steps were made, and in which order, during the simulation process.

In chapter five, the results of the research will be presented with the help of clear and illustrative graphs and tables. Firstly, temperature profiles of each temperature region will be presented, showing how the temperature during the simulations fluctuate around the target value. After that the results regarding local atomic structures in three different temperature regimes will be presented including the radial distribution functions, coordination numbers and bond angle distributions. Lastly, in chapter five, results regarding atomic mobility in three different temperature regimes will be presented including the mean square displacements and diffusion coefficients.

In chapter six, conclusions about structural geometry, comparisons on different environments and the obtained information and calculated values will be discussed, and how they align with previously performed research on the same element composition.

In chapter seven, the outlook for future research will be discussed including what kind of other properties could be studied from the gathered data from this thesis, more temperatures areas to be studied and further analysis.

2 Phase change materials

Phase change materials are materials which are able to at least exist in two structurally different phases of solids, amorphous and crystalline. These amorphous and crystalline phases have many differences in electrical and optical properties that are caused by the differences in their structures. These differences in the structures and their properties make it possible to use phase change materials in storage applications and make the information storing possible. [2]

Phase change materials, based on chalcogenide alloys, have gained much interest because of their ability to exhibit fast and reversible transitions between two phases, crystalline and amorphous, due to Joule heating [6]. Phase change memory materials are next-generation contenders for the technology of non-volatile electronic-memory, potentially replacing semiconductor flash memories made from silicon complementary metal-oxide [7].

2.1 Amorphous phase

The amorphous (or glass) phase refers to atomic structure that is disordered, meaning it lacks clear long-range atomical arrangement, as can be seen in figure 1 [1,2,8]. Due to the lack of periodicity and long-range atomical arrangement, a unit cell or its atomic positions in the unit cell cannot be defined. In addition, the precise identification of structural properties of amorphous materials corresponds to an arduous task. However, certain statistical properties, such as the correlation function and the structure factor, for instance, can be determined successfully to classify and characterize the amorphous structure by observing the short and medium range order. [2]

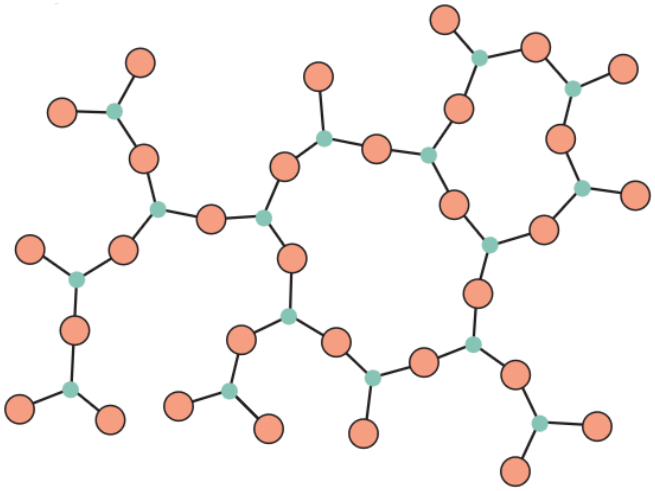

Figure 1 General illustration of amorphous phase. The disordered structure of silicon dioxide (SiO_2). Green color presenting silicon and orange color presenting oxygen atoms. Reproduced with permission from John Wiley & Sons [9].

2.2 Crystalline phase

The crystalline phase, unlike the amorphous phase, has ordered repeating long-range atomical arrangement, as can be seen in figure 2. Due to the repeating long-range atomical order crystalline phase has unit cells and its atomic positions can be defined accordingly. [9] Even though the crystalline phase has an ordered structure, there can still exist significant number of vacancies in the structure [1].

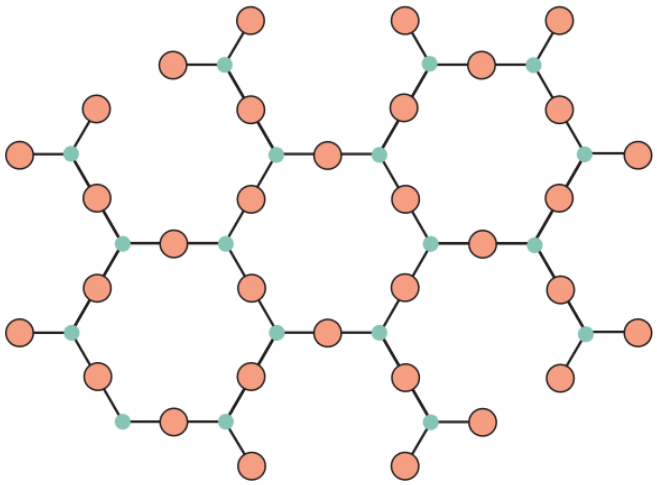

Figure 2 General illustration of crystalline phase. The ordered structure of silicon dioxide (SiO_2). Green color presenting silicon and orange color presenting oxygen atoms. Reproduced with permission from John Wiley & Sons [9].

2.3 Electronic memory and digital storage

Phase change materials applications in solid state memory devices and optical storage depend on the properties of the materials [2]. Crucial properties for phase chance materials are the structure of the amorphous and crystalline phases [1], and the fast switching between these two phases repeatedly over many cycles from 10^5 to 10^{17} times depending on the application. The crystallization time of the material determines the data rate inside the memory device and the target is to be short. For optical storage applications, to cause phase transition, the given wavelength of laser light has to be adsorbed appropriately. This is also a reason that the most successful phase change materials are based on chalcogenide alloys. For phase change random access memory applications, it is important that the material has high resistance contrast between the two structural phases. [3] Many phase change materials have contrast in the resistance between crystalline and amorphous phases [1,3], with values ranging from four to five orders of magnitude. During the melt-quenching process, the low melting temperature of chalcogenides is advantageous, as well as the low thermal conductivity, allowing the size of phase change random access memory chips to be minimized. [3]

In optical storage applications, phase change materials have been used in the first-generation rewritable compact disks, CDs, the second-generation rewritable digital versatile discs, known as DVDs, and in the third-generation Blu-ray discs [1–3]. The storage capacity in each generation has been steadily increasing, from early CDs being 500 MB to today's dual layer Blu-ray discs being 50 GB. This is made possible because of multiple different factors such as higher numerical aperture, shorter wavelength and smaller size of written bits that lead to an increase in storage density. [3] In non-volatile electronic data storage applications, phase change alloys have been intensely discussed to be suitable material. Phase change materials have much potential for future applications and therefore they are investigated actively by industrial semiconductor companies. [1]

2.4 Phase change process

In phase change materials applications, the crystallization process time is the key data-rate-limiting process and therefore this parameter is very important. The speed of crystallization process determines the data transfer rate. [2,3] The main reason for the crystallization to take place is the Gibbs free energy difference between the amorphous phase and crystalline phase

[2]. In the crystallization process, which is also called the SET process, the material can be switched from glass to crystal. In the amorphization process, that is also called the RESET process, the crystal is switched back to glass. [3] The whole phase change process can be performed repeatedly [2,3]. The phase change process that includes both crystallization and amorphization processes is illustrated in figure 3. In the SET process a low pulse of voltage is applied, and the crystallization of the amorphous phase takes place, that enables the data to be stored. In the RESET process, a higher pulse of voltage is applied with lower duration leading to the amorphization of the crystalline phase. Typically, the phase-change process takes place in nanoseconds. The binary bits of information, "0"-bit and "1-bit", are stored as two different structural states of the same material, glass (high resistance) and crystal (low resistance), respectively. [3]

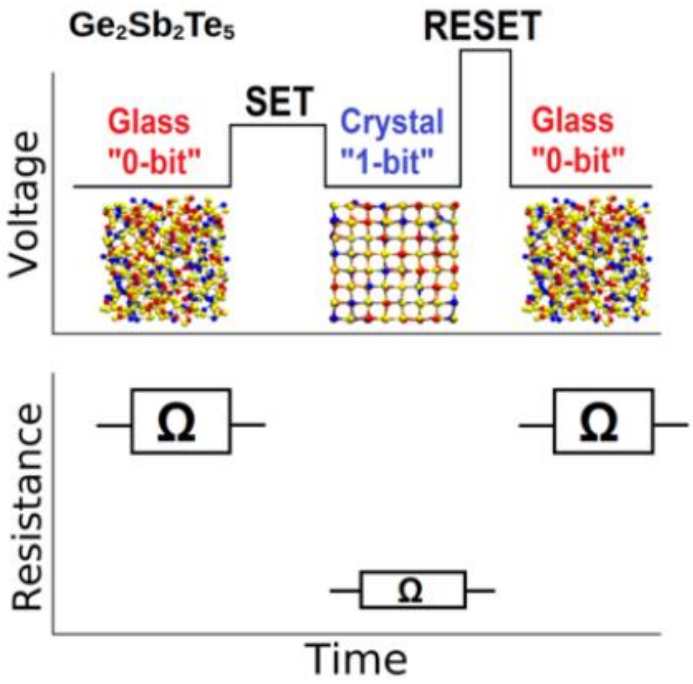

Figure 3 Illustrating the phase change process and the digital storage of the information, including the crystallization process called as SET, and amorphization process called as RESET.

2.5 Fragility

Fragility is a concept for describing the temperature dependence of viscosity [10,11]. Fragility is central to understanding the supercooled liquid state of phase change materials, and is account for the temperature dependence of the atomic mobility, which is relevantly connected to phase switching kinetics [11]. Commonly, fragility can be characterized by measuring Angel plot´s

slope at glass transition temperature. In this concept, liquids can be classified into two groups, strong and fragile, based on the behavior of viscosity as a function of temperature. Liquids with viscosity that shows near-Arrhenius rise when approaching glass transition temperature while cooling are thought of as strong liquids, and liquids that show non-Arrhenius behavior thought to be fragile liquids. In general, phase change materials are fragile systems which roots to their poor glass forming ability and fast crystallization speed. [10,11]

3 Methods

To be able to perform computational studies efficiently, correct and efficient methods have to be chosen. In the process of selecting appropriate methods, their limitations must be also understood. The right choice of methods makes studies and research work possible to be carried out effectively.

3.1 Computer simulations

During the last decades, computational studies have been gaining popularity because of the increase in the computer power and the development of efficient numerical algorithms. The computer simulations can be considered as "*in silico*" type of experiments that operate as a bridge in between theories and real experiments. Figure 4 illustrates the connection between computer simulations and experiments. In addition, they are capable of connecting the microscopic to macroscopic behavior of the system. Computer simulations are acknowledged as effective tools in studying complex systems, relatively inexpensive and simple to perform. In principle, everything in a simulation can be measured, allowing us to gain a view about the materials properties. [12]

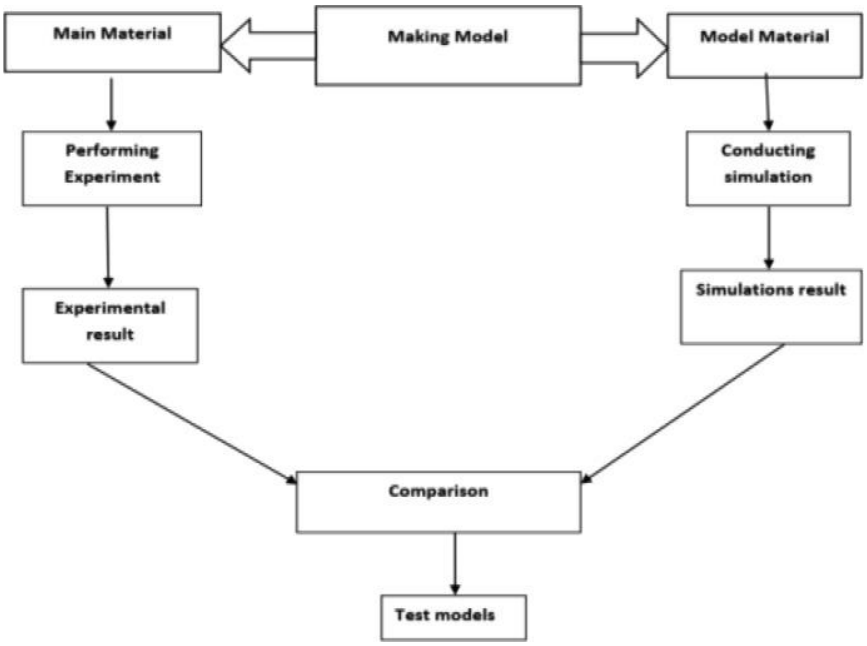

Figure 4 The connection between computer simulations and experiments. Reproduced with permission from Elsevier [13].

With the help of computer simulations, the detailed level of a system can be investigated in a way that would not be possible in analytical theory or in real life experiments. On the one hand,

essential information from the atomistic level will be gained from the simulations that are useful to rationalize and to tailor processes, which could be done experimentally or propose new experimental pathways. On the other hand, simulations are able to solve problems by a numerical method, which can give the possibility to improve and validate theoretical models that already exist. Also, through a simulation, particularly important macroscopic properties are possible to be investigated and studied via microscopic modelling, which would not be attainable with experimental ways or otherwise it would require really expensive experiments. [12]

During the early days of computer simulations, they were mostly used in problems related to statistical physics addressing different questions. In the recent era that high performance computing facilities have been developed, and continuously becoming more powerful and accessible, computer simulations became an important tool for studying real materials properties. The two main types of computer simulations to study material properties are ab initio and classical approaches methods. The most common method in the classical approach is molecular-dynamics simulations. [12]

3.2 Classical molecular dynamics

A traditional computational method for material modeling is classical molecular dynamics. In classical molecular dynamics the effect of both time and temperature are included. Classical molecular-dynamics simulations have achieved significant success in a wide range of modelled systems, in probing the dynamics and the atomic structure. The application of classical molecular dynamics in studies allows the possibility of atomic scale characterization of the structure, the atomic structure composition relationships investigation, providing as well as insights into atomic migration mechanisms in materials. The principle behind classical molecular-dynamics simulations is based on the evolution of the modeled atomic system that is followed as a function of time by generating a continuous trajectory in phase space. [12]

3.2.1 Equations of motion

In classical molecular dynamics the goal is to define the trajectory of the particles in the modeled system. In order to be able to do this, the equation of motion will be used. The equation of motion can be presented in three different forms: Hamiltonian, Newtonian and Langrangian, which all correspond to each other. Using the Hamiltonian equation representation, the

velocities and positions of the systems every atom are defined from classical mechanics with using the equation of Newton´s second law [12]:

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i \quad (1)$$

where m_i presents the mass of a certain atom *i*, the $\vec{r}_i$ the atoms position i, and the $\vec{F}_i$ the force that acts the atom. Considering an interacting set of particles, *N*, which are in a conservative field moving. The forces that act on the set of particles can be defined with the help of the function of potential energy, *V*, and with that the equation of motion are able to be modified and put as follows [12]:

$$m_i \frac{d^2\vec{r}_i}{dt^2} = -\nabla_{\vec{r}_i} V(\{\vec{r}_i\}) \quad (2)$$

Here, the $\{\vec{r}_i\} = (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ refers to the total number 3*N* coordinates of complete set of particles. This leads to the need to solve a set of 3*N* coupled second-order ordinary differential equations in order to gain the atoms trajectories. In principle, physical properties are being able to be calculated from the collection of *N* particles based on the knowledge of phase space trajectory of the particles. The first order derivative and the force are both needed by the integration in order to be continual, which then indicates that the function of the potential energy. This means that the first and second order derivatives of it shall also be continuous. Additionally, the integration constants of 6*N* specification are needed, meaning 3*N* for the positions and 3*N* for the velocities of the particles. This corresponds to the initial conditions of the system at time being zero, $\{\vec{r}_i(0), \dot{\vec{r}}_i(0)\}_{i=1,2,\ldots,N}$. From the available structure data and crystallographic tables (information) from the literature, the initial positions $\{\vec{r}_i(0)\}$ are able to be defined, that are usually gathered from the diffraction pattern. The starting velocities, $\{\dot{\vec{r}}_i(0)\}$, are possible to be chosen randomly or optionally to meet the distribution of the Maxwell-Boltzmann that corresponds to wanted temperature. An alternative approach to this is to begin with particle velocities being zero and then rescale to the point that the target temperature is reached and maintaining this temperature. [12]

3.2.2 Schemes of integration

A certain atom´s potential energy depends on its interactions with the system´s other atoms. Because every particle is interacting with all the other particles in the system, a body problem of $N - 1$ is needed to be solved, meaning the equation 2 cannot be solved in an analytical way. Instead, by using numerical methods for the integration, coupled set of 2nd order differential equations can be solved in an approximate way. [12] One perspective to view the numerical methods is to think that numerically computed trajectories are close to the true trajectories of the physical problem [14].

When time is treated as a discrete quantity, $t = k\Delta t$, and with using the initial conditions knowledge, the equations motion for the $3N$ set are able to be solved to velocities, and yield positions at some times later, $t + \Delta t, \{\vec{r}_i(t + k\Delta t),\ \dot{\vec{r}}_i\,(t + k\Delta t)\}_{k=1,2\ldots,L}$. By applying this way of approach, in the phase space a discrete trajectory is generated over consecutive time intervals, Δt, that are called as a timestep. The numerical integration scheme´s success is significantly affected by the chosen value of the timestep. The larger the value of Δt the less number of integration steps there will be over certain amount of simulation time there will be, reflecting less computational cost. However, this means that the larger the Δt is the poorer the trajectory sampling in the phase space will be, and the solution will be less accurate for the equation of motion to approximate the trajectories correctly. To put contrast, when the timestep is smaller the more accurate that is, and more exact the sampled trajectories and the solution of the equation of motion will be. The aim is to balance between computational expense and the accuracy of the computational results. [12]

A numerical algorithm or in other words an integrator is needed in order to propagate the solution and thereby the trajectories of the particles step-by-step. In order to receive stable trajectories even by using a relatively large timestep, Δt, an efficient and well behaving integrator that is simple and fast enough has to be used. Additionally, the mechanical energy has to be preserved and the temporal evolution has to be reversible. According to Liouville's theorem, the forces are needed to be computed by the integrator, but not too often, and the volume of the phase space must be maintained. [12] A variety of different numerical methods to integrate the equation of motion are available [14]. All the different integrators are based on Taylor's expansion in time, t, in order to receive velocities and positions at some time later, $t + \Delta t$ [12].

Integrators have been traditionally classified according to their accuracy and stability. Normally, accuracy refers to precision with respect to the trajectory that is given, and stability refers to boundedness with respect to interferences. Both of these concepts have limited practical value on the simulation duration, as well as on the treatment of nonlinear dynamics, because of the sensitive dependence on initial conditions. Nevertheless, it is considered that accuracy and stability are basic prerequisites for an effective integrator in order to perform sufficiently in molecular-dynamics simulations. [15]

The Verlet algorithm is one of many numerical integration schemes that is simple to implement. Verlet algorithm integrator has many different advantages that include, for instance, time-reversibility and great stability. In order to start the Verlet algorithm, the Taylor's expansion has to be considered for two different timesteps, $\pm\Delta t$ with the position vector at some time, *t,* as follows [12]:

$$\vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \frac{d\vec{r}_i(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}_i(t)}{dt^2}\Delta t^2 + \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \qquad (3)$$

$$\vec{r}_i(t-\Delta t) = \vec{r}_i(t) - \frac{d\vec{r}_i(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}_i(t)}{dt^2}\Delta t^2 - \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + O(\Delta t^4) \qquad (4)$$

These two expansions presented above summed up together will result for the expression of the Verlet algorithm, as follows [12]:

$$\vec{r}_i(t+\Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t-\Delta t) + \vec{a}_i(t)\Delta t^2 + O(\Delta t^4) \qquad (5)$$

The acceleration, $\vec{a}_i(t)$, which is presented in the equation is obtained by calculating the corresponding force that is defined from the position $\vec{r}_i(t)$ in equation 2. Hence, this summed up scheme of integration only stores up the forces and the positions. The errors in the Verlet algorithm are of 4th order, which can be thought to be very small when compared to the intermolecular distances. The Verlet integrator´s disadvantage is mainly that it is missing a velocity term which leads to a situation that estimating the particles velocities is not a straightforward and therefore requires the determination of the position vector first at $t + \Delta t$. After that the two Taylor's expansions need to be deduced from equations 3 and 4. Thereby, the velocities are indirectly computed through finite differences as follows [12]:

$$\vec{u}_i(t) = \frac{\vec{r}_i(t+\Delta t) - \vec{r}_i(t-\Delta t)}{2\Delta t} + O(\Delta t^2) \qquad (6)$$

The obstacle related to the calculation of the positions at $t + \Delta t$ with the Verlet integration scheme is that the knowledge of the positions at the previous step, $t - \Delta t$, and at the time, *t,* are needed both, and the Verlet integration scheme is not a type of self-starting algorithm. Additionally, the calculation errors in the velocities are of 2nd order, leading to situations where the calculations of the positions are significantly less accurate. To overcome the need of storing the frame data from the previous simulation to calculate velocities with Verlet integration algorithm, the velocity Verlet integration algorithm is used, which is typical´s Verlet algorithm´s extension. The calculations of the position and velocity of every atom in the velocity Verlet integration algorithm are done as follows [12]:

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{u}_i(t)\Delta t + \frac{1}{2}\vec{a}_i(t)\Delta t^2 + O(\Delta t^4) \qquad (7)$$

$$\vec{u}_i(t + \Delta t) = \vec{u}_i(t) + \frac{1}{2}[\vec{a}_i(t) + \vec{a}_i(t + \Delta t)]\Delta t + O(\Delta t^3) \qquad (8)$$

By using the equation 8 the velocities can be calculated accordingly. The velocity Verlet algorithm can be seen as predictor-corrector, in which the atoms positions and the velocities of in some time in future are able to be calculated based on the positions and velocity knowledge of the atoms at a certain time. The systems state in the future is able to be calculated by repeating this procedure until reaching the target time. In materials modelling studies one of the most broadly used is the velocity Verlet algorithm, since it is able to fulfill the necessary requirements of an efficient integrator, such as appreciable numerical accuracy and feasible implementation. [12]

3.2.3 Interatomic potential

In classical molecular dynamics, the particles, which construct the material, interact with each other, and these interactions are described with an interatomic potential [14]. Since the interatomic potential describes the forces between atoms in the modelled system, it plays a significant role in the quality and accuracy of the calculated results of the simulation. The intermolecular force fields are conservative. Because of this the potential energy function can

be used to calculate the potential energy of a system. The intermolecular force fields define different types of interactions between atoms and describe the system´s potential energy that contains *N* number of atoms depending on the atoms coordinates, $\{\vec{r}_i\}_{i=1,2,\dots,N}$. The systems total potential energy of *N* number of atoms with the many-body expansion can be expressed as follows [12]:

$$V_{tot}(\{\vec{r}_i\}) = \sum_i V_1(\vec{r}_i) + \sum_i \sum_{j>i} V_2(\vec{r}_i, \vec{r}_j) + \sum_i \sum_{j>i} \sum_{k>j} V_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) + \dots \qquad (9)$$

where all of the different pairs between *i* and *j* are summed up without counting any of the pairs more than once, and $\{\vec{r}_i\} = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ corresponds to the complete set of the *3N* particle coordinates. $V_1(\vec{r}_i) = V_i$, which is the first term, represents the energy of the single particle that in normal situations is equal to zero, with exceptions related to cases where the particles are affected by an external force field. $V_2(\vec{r}_i, \vec{r}_j) = V_{ij}$, that is the second term, represents the two-body interaction, where the interaction between two atoms depends only on the distance between the two atoms. $V_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) = V_{ijk}$, which is the third term, represents three-body interactions, and describes the change in the two-body interaction due to the presence of the third atom. [12]

An accurate two-body interaction, from the equation´s 2 point of view, is considered as a significant input since it is needed to compute accurate trajectories of the particles in the phase space. The two-body interaction can be described by a force field, which in turn is defined by a set of analytical functions and specific parameters that are depending on the particle's relevant positions. Therefore, the system´s total potential energy can be calculated by the sum of the two-body contributions to the energy as follows [12]:

$$V_{tot} = \sum_i \sum_{j>i} V_{ij}(r) \qquad (10)$$

where $r = |\vec{r}_i - \vec{r}_j|$, represents the separation of the pair, and each atom pair interaction depends only on the distance of this. The potential energy of a pair of species that interact are able to be distinguished to long- and short-range contributions as follows [12]:

$$V_{ij}(r) = V_{long}(r_{ij}) + V_{short}(r_{ij}) \qquad (11)$$

The long-range contribution, also called as Coulombic term, arises because of the attraction/repulsion electrostatics that occur between the ions. Usually, the Ewald summation is used as a method to describe these interactions in a molecular-dynamics simulations. The short-range term can be defined as a Van der Waals interaction that occurs between two ions. At small interatomic distances, the short-range contribution is dominant, and therefore it determines the behavior of the atom. After a certain cut-off distance, the short-range interaction loses its effect, and this needs to be defined in the calculation. The short-range interactions need to be described in a way that they are consistent with computational efficiency and energy conservation. In a case where two particles approach each other closely, the electron cloud overlap may occur, which then can give an increase to repulsive force. This phenomenon arises from the Pauli´s exclusion principle. Thus, the repulsive potential energy increase of Pauli must be taken in to account in the potential model´s total expression, which then leads to a better result on description of the forces. [12]

The model of interatomic potential is typically parametrized from experimental data by fitting the structural, and in some cases the elastic and vibrational, properties, which correspond to specific conditions of pressure and temperature. The two-body potential describes usually suitably enough the expression of the potential energy surface at the minimum, that is referred as the equilibrium separation. This is thought to be rather specific to each modelled system, meaning that it demonstrates only small transferability under configurational changes and/or different physical conditions of the variations. Carrying out classical molecular-dynamics simulations with accurate interatomic potentials is computationally inexpensive, and it gives the opportunity to model large sized systems with long timescale sampled trajectories. [12]

3.2.4 Periodic boundary conditions

Physical system sizes in the size 10^{23} atoms are practically out of the reach computationally. Molecular-dynamics simulations are able to model relatively small systems of about 100 – 1 000 000 particles of “real” material, to obtain reasonable results that are in good agreement with experimental data [16]. In a molecular-dynamics simulation the system which is modelled is placed into a finite box with a specific geometry. This is often called as simulation cell. The simulation cell´s specifications of in size and shape, as well as the total atom count are connected to the density and molar composition of the modelled system. The relatively “small” system size raises several challenges associated with its ability to sufficiently describe the properties of a macroscopic sample. An example of this is that the volume element is

constrained in a material by the surrounding bulk, and such constraints break down at the surface leading to extensive atomic relaxation that can often occur. Systems which are macroscopic contains only a little proportion of the atoms in the surface, and thus contributes only slightly to bulk properties. A different situation is related to cases where the simulation cell is constituted by a small number of atoms, which leads to a different surface area to volume ratio. This means that the environment is different for the few atoms in the surface compared to the atoms in the bulk. Therefore, the surface atoms will experience different forces compared the bulk region atoms, which could affect the modelled system´s properties, and might prevent the attempts to obtain the properties of bulk material from such a system. [12]

In order to avoid the contribution of the surface effects and to overcome the obstacles caused by the finite size of the simulation cell in a molecular-dynamics simulation, periodic boundary conditions are applied to make infinite bulk around the sample. This can be achieved by replicating the simulation cell in *x*, *y* and *z*-directions as shown in figure 5. [12]

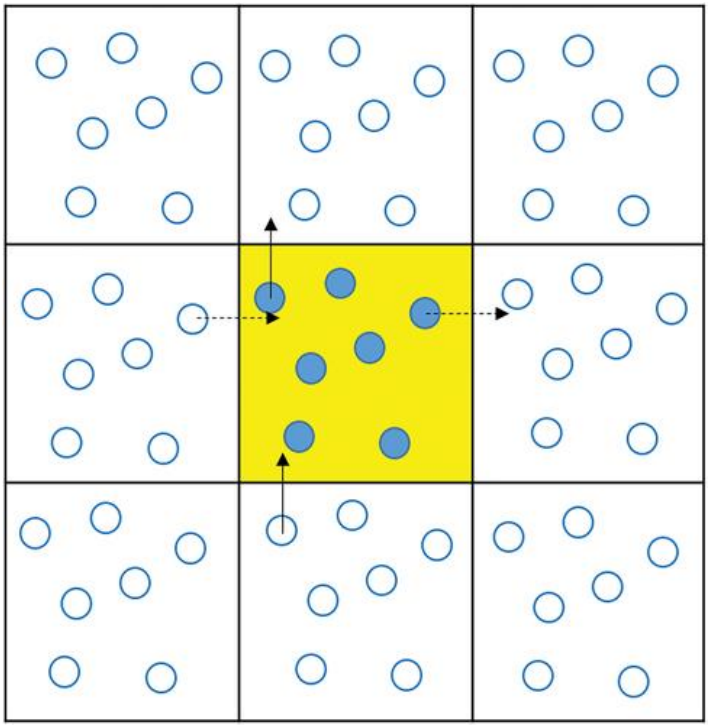

Figure 5 A two-dimensional illustration of periodic boundary conditions of the simulation cell replications in x and y dimensions. Reproduced with permission from John Wiley and Sons [17].

Since the replication of the simulation cell in all directions is associated with the modelled system and to the total number of atoms in it, the atoms motion in the initial simulation box is correlated with their motion to the replicated simulation cell. This means that when an atom

leaves from one side of the initial simulation box, it enters back to from the opposite side of the box, while having the exact same velocity. In other words, when any of the atoms crosses the periodic boundary, it appears to the simulation cell from the other side. Hence, according to this framework of periodic boundary conditions, the simulation cell has no wall and no atoms in the surface, while inside the simulation cell´s the total atom amount is kept the same. In molecular-dynamics simulations periodic boundary conditions are applied in order to model crystals, liquids or glasses. Additionally, the shape of the simulation cell might vary due to the characteristics of the modelled system. [12]

3.2.5 Ensembles

The positions and velocities of the atoms of the modelled system are the output of the numerical integration of equation 1. The bulk and macroscopic properties of the system have to be harmonized with the atomistic quantities that are then examined by the simulation, such as the temperature, volume and pressure of the simulation. The bridge between the macroscopic properties and the atomic-scale properties is provided by statistical mechanics. The thermodynamical ensembles correspond to the core idea of this concept. [12]

During the molecular-dynamics simulation, the macroscopic properties are kept constant, which can be characterized by an ensemble. The modelled system´s total energy, *E,* as in the Newton's equation of motion, is kept constant during the integration for *N* number of particles in fixed volume of the cell, *V*. Consequently, the system´s state at each molecular dynamics timestep is a representative of an ensemble where extensive variables are the number of atoms, *N*, the volume of the simulation cell, *V*, and the total energy, *E*. This represents the microcanonical ensemble, *NVE*, which is considered as the natural ensemble in molecular-dynamics simulations. It corresponds to a computational method to propagate a system following the constant energy´s path in the phase space, and due to this the true dynamics of the simulation can correctly be reproduced. Practically, in molecular-dynamics simulations for equilibration runs the *NVE* ensemble is used until reaching the target, average temperature of the system. The *NVE* ensemble, for production runs, is used to collect samples of the trajectory that will be then further analyzed to compute the properties which are though relevant. [12]

During the molecular-dynamics simulations, the temperature needs to be controlled and for this purpose a thermostat is used. It fixes the temperature to a specific value and during the simulation acts as a thermal energy source, supplying or removing heat accordingly. This

corresponds to the working principle of the canonical ensemble, which allows the temperature to vary in the simulation, while taking care that the temperature remains constantly averaged in a time fashion. The canonical ensemble corresponds to the thermodynamical state of the collection of systems which are characterized by fixed amount of atoms *N,* fixed volume *V*, at fixed temperature *T*, typically referred as *NVT.* One of the available thermostats for molecular-dynamics simulations is the Nosé-Hoover thermostat. It uses a friction coefficient in order to describe the tailored equations of motion and control the variations of the temperature. The Nosé-Hoover thermostat generates strictly the thermodynamics of the canonical ensemble and is able to approximate the true dynamics of the simulated system, due to the fact that the time evolution is deterministic and stochastic changes are not included. However, it should be noted that the Nosé-Hoover thermostat can sometimes exhibit pathological behavior in some specific modelled systems. [12]

During the molecular-dynamics simulations, often comes a case study where both the temperature and pressure are desired to be maintained constant. In this kind of situation, the simulation is carried out by using an isobaric-isothermal ensemble, that is characterized by a fixed number of atoms, *N*, fixed pressure, *P*, and fixed temperature, *T*, typically referred as *NPT*. Utilizing this ensemble structural modifications due to phase transition or pressure can be studied. In practice, this approach is the closest and most realistic to experimental procedures. From a thermodynamical point of view, the volume becomes a dynamic variable, and when the pressure is constant the system is able to change volume with their own surroundings. Therefore, a change in cell volume is allowed and due to this, volume fluctuations are involved at constant pressure in a simulated system. There are two *NPT* ensemble families, the isotropic and anisotropic. In isotropic *NPT* ensemble, the volume is allowed to change but the simulation box´s shape is kept the exact same. In anisotropic *NPT* ensemble, the volume as well as the simulation box´s symmetry are both allowed to change. Performing a molecular-dynamics simulation with the *NPT* ensemble, the thermostat is combined with a barostat in order to control and keep the pressure of the system close to the desired value. A number of different barostat techniques are available, and they work by adjusting the volume of the simulation cell. It is noted that many of these barostats are counterparts to the respective thermostats. [12]

3.2.6 LAMMPS

LAMMPS is an open-source code that is used to perform classical molecular-dynamics simulations, focusing on materials modelling. The abbreviation LAMMPS comes from the

words Large-scale Atomic/Molecular Massively Parallel Simulator. LAMMPS has several available potentials for many different materials, allowing the simulations of soft matter, solid-state materials, and mesoscopic and coarse-grained systems. Also, LAMMPS can be used to model atoms or generically different scale systems, such as atomic, meso or continuum as parallel particle simulator. [18] The major advantages of LAMMPS are related to its extensive features list, the extendibility of its codebase, and the highly scalable performance [19]. LAMMPS is able to run on a single processor or have the capacity to run in parallel using message-passing techniques and simulation domain spatial-decomposition [18].

3.3 Machine-learning potentials

Currently, different areas of chemistry, physics and materials science are experiencing a tremendous impact from machine learning methods, where different research questions are ranging from the accurate molecular energies prediction and molecular reaction mechanism exploration all the way to screening of functional materials through a target property optimizing as a function of different descriptors. [20]

A machine-learning potential represents the $3N$-dimensional potential-energy surface in a mathematical form. Meaning, that it is the total interatomic forces and energy for a certain set of N atomic positions. The concept resembles force fields that are fitted empirically, and machine-learned potentials make assumptions, physically motivated, such as smoothness and locality of the potential-energy surface. The key difference between traditional and machine learned potentials is that machine learned models don't make any prior assumptions about the interatomic potential´s shape, i.e., the specific functional form of the potential-energy surface as a function of the atomic positions. Instead of that, from a large set of input data, all the relevant information is directly extracted. Then the data are computed with reference level, which is much more accurate and computationally expensive. After fitting the potential, it can be applied to predict forces and energies for larger atom ensembles without additional reference data needed. By that, it allows the users to perform molecular-dynamics simulations that are faster with orders of magnitude from the reference method. Because of this, more complex problems regarding atomic structure can be brought to computational reach. Table 1 presents some examples of methods and implementations currently available for machine learned interatomic potentials. [20]

Method	Refs.		Regressor	Implementation
Artificial neural networks (Behler)	40	(2007)	NN	Standalone ("RuNNer"); LAMMPS interface67
Gaussian approximation potentials (Bartók and Csányi)	68	(2010)	GPR	GAP code (custom); LAMMPS interface
Spectral neighbor analysis potential (SNAP) (Thompson)	55, 69	(2015)	Linear fit	LAMMPS interface
Adaptive, generalizable, and neighborhood informed (AGNI) force fields (Ramprasad)	70-72	(2015)	KRR	LAMMPS interface
aenet (Artrith)	73	(2016)	NN	Standalone ("aenet")
Amp (Korshidi and Peterson)	74	(2016)	NN	Standalone ("amp"); LAMMPS interface
Moment tensor potentials (Shapeev)	54, 75	(2016)	Linear fit	LAMMPS interface
DeePMD (E)	76	(2018)	NN	Standalone ("DeePMD-kit"); LAMMPS interface

Table 1 Examples of methods and implementations related to machine learned interatomic potentials. Reproduced with permission from John Wiley and Sons [20].

3.3.1 Types of different machine learning interatomic potentials

Currently many different machine-learning interatomic potentials exist for different applications to increase accuracy and computational efficiency in atomistic simulations. They are used widely in different fields, varying in different materials and their applications such as ceramics, metals and alloys. Examples of machine-learning interatomic potentials used in materials science studies include high-dimensional neural networks (HDNN), Gaussian approximation potential (GAP), moment tensor potential (MTP), deep potential molecular dynamics (DeePMD), and message passing neural network (MPNN). [21]

3.3.2 Gaussian Approximation Potential (GAP)

Gaussian approximation potential (GAP) is a machine-learned potential that follows Gaussian process regression. GAP is part of a class called Kernel methods and is reliably used in wide range of applications [22]. One reason for the wide range of its applications is the advantage of

GAP to be very flexible. Gaussian process uses forces and energy for regression and its atomic environment representation is complete. [23,24]

3.4 Modelling properties

In this thesis, different properties have been studied that can be classified into two different categories: short-range order structural properties and ionic diffusion properties. Short-range order structure describes how the atoms are arranged in the structure of the material in their first nearest-neighbour distances. Ionic diffusion properties describe the atomic mobility inside the structure of the material. In this thesis the short-range order structural properties that were studied are radial distribution function, coordination number and bond angle distribution. The ionic diffusion properties that were studied are mean squared displacement and diffusion coefficient.

3.4.1 Short-range order structure

The radial distribution function, also called pair correlation function or pair distribution, *g(r)*, is a structural property that can be calculated from a molecular-dynamics simulation [12]. It describes how atoms are arranged in the structure around a given atom [12], presenting the probability of finding an atom in a shell, *dr*, at certain distance, *r*, from a chosen another atom as a refence point, as presented in figure 6 [12,25].

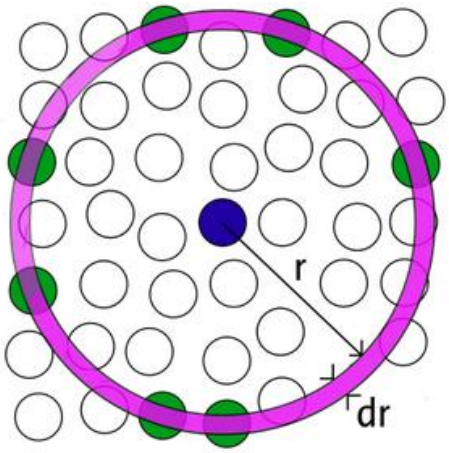

Figure 6 Radial distribution function, g(r), illustrated in two-dimensional space [25].

The coordination number is a structural property, which describes the number of closest other atoms that are bonded into a specific atom, given a specific geometric cut-off distance [26]. The

coordination numbers are used to describe the local atomic environments inside a model structure [12]. The figure 7 illustrates the definition of coordination number.

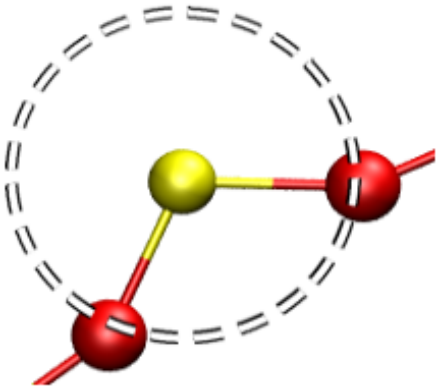

Figure 7 Illustration of coordination number and how it is determined [27].

The bond angle distribution, $g(\theta)$, is a structural property that gives additional information on the local geometry of the first coordination shell of each atomic species inside the material structure [12]. It provides the measure of the probability that a central atom and two of its neighbor atoms form an angle to directions away from the central atom, for instance, A–B–A, with atom B as the central atom [28]. From the bond angle distribution, one can gather information about the different kinds of molecular geometries the atoms can adopt within the simulated structure [12]. The figure 8 illustrates the definition of bond angle.

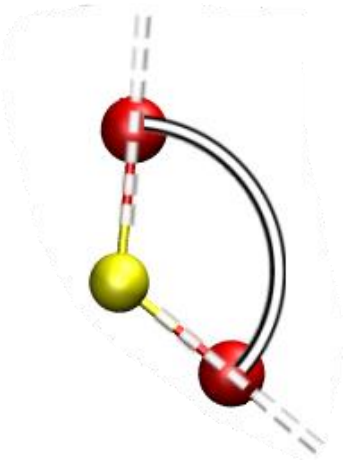

Figure 8 Illustration of bond angle and how it is determined [27].

3.4.2 Ionic diffusion

The mean squared displacement (*MSD*) is a dynamical property that describes the mobility of the different atoms in the structure [29]. Mean squared displacement is a quantity that can be measured, and is directly linked to the motion of atoms. It measures the average distance of a certain atom travelled in time, and it can be defined by the following equation [12]:

$$MSD(t) = \frac{1}{N}\sum_{i=1}^{N}\langle|\vec{r}_i(t) - \vec{r}_i(t=0)|^2\rangle \qquad (12)$$

Where the vector $\vec{r}_i(t) - \vec{r}_i(t=0)$ is the distance that the atom *i,* has travelled by some intervals of time, *t*. The vector´s squared magnitude is averaged over many time intervals, and also over all the same atom species in the structure. [12]

The diffusion coefficient, *D*, is a dynamical property that can be calculated for a specific particle type, *i* [12]. The diffusion coefficient describes the rate of transfer of particles, which are diffusing across the given unit area [30]. By using the calculated mean squared displacement and through the Einstein relation equation, the diffusion coefficient can be determined as follows [12]:

$$D_i = \frac{1}{6}\lim_{t\to\infty}\frac{d}{dt}MSD(t) \qquad (13)$$

In this scenario, the diffusion coefficient can be obtained via the slope of the linear regime of the mean squared displacement that is plotted as a function of the simulation time. Usually, it is found that temperature has a significant effect on the calculation of the diffusion coefficient. [12]

4 Computational details

The simulations were performed with using a 900-atom model of the ternary $Ge_2Sb_2Te_5$ composition in three different temperature regions: overheated amorphous, supercooled liquid and liquid. These three relevant temperatures were 420 K for overheated amorphous, 825 K for supercooled liquid, and 1200 K for liquid. The amorphous $Ge_2Sb_2Te_5$ models were generated with the melt-and-quench method by using molecular-dynamics simulations with a machine learned Gaussian Approximation Potential. The number of atoms was chosen to be 900 in order to balance between accuracy and computational cost associated with the molecular-dynamics simulations. The total simulation time was set to be 400 picoseconds (ps) for each molecular-dynamics simulation, at each temperature, and the timestep was set to be 2 femtoseconds (fs). The simulation cell was 30.722 Å in *x*, *y* and *z* directions with 90 degree angles for all of the edges, corresponding to a cubic simulation box.

4.1 Input file

The input file is one of the two files needed for the calculations using LAMMPS. The input file tells the simulations all the commands needed to run the molecular-dynamics simulation, which the high-performance computer reads line by line executing them in the same order that the command lines are written in the file. Figure 9 shows the settings that were used in the simulations, which tell the simulation all the essential information. Command lines with hashtag marks refer to comments and they are not taken into account in the execution. The "units" command tells what kind of specified units are used in the simulation, and "metal" referring to certain set of units for the modelled system for calculating the properties [18]. The "boundary" command tells in what kind of style the simulation box dimensions are treated as, and the "p p p" refers to periodic style for all *x*, *y* and *z* directions, hence enforcing periodic boundary conditions in all directions. The "read_data" command tells the simulation the data file where information regarding the initial model structure is read, and "GST_data" is referring to this data file. The "read_restart" command tells the simulation the previously saved configuration that will be read to continue the simulation from the previous run (i.e., from the last saved molecular dynamics step), and the "rst.23500.gst" refers to this data file. The "pair_style" command tells the simulation the type of the pairwise interactions that are used, and "quip" refers to this pairwise interaction, associated with the machine-learned GAP potential. [18] The "pair_coeff" command tells the simulation the pairwise force field coefficients that are used for the atom types (through the atomic numbers of the elements) [18],

and “gp_merged_2bdmbd.xml” refers to this file, which is the specific version of the GAP potential used in this study [31]. The “neighbor” command tells the simulation the parameters that has affect to pairwise neighbor list building, and “0.3 bin” refers to the cut-off distance in Å. The “neigh_modify” command tells the simulation the modified parameters to be used for the pairwise neighbor lists, and the “delay 10” refers to how many steps the simulation should wait until new list is made. The “thermo_style_custom” command tells the simulation in what kind of customized style the thermodynamical data of choice should be printed out. The “thermo” command tells the simulation how often the thermodynamical data should be printed out, and the “50” refers to the number of timesteps, i.e., how frequently the data are printed in the output file of the simulation. [18]

```
##########################################################
#Amorphous GST using GAP
##########################################################

units           metal

boundary        p p p

atom_style      atomic

read_data       GST_data
read_restart    rst.235000.gst

pair_style      quip
pair_coeff      * * gp_merged_2bdmbd.xml "IP GAP" 32 51 52

neighbor        0.3 bin
neigh_modify    delay 10

thermo_style custom step temp etotal pe ke evdwl press vol enthalpy
thermo          50
```

Figure 9 Molecular-dynamics simulation settings that were used in this study.

The melt-and-quench process for the generation of the atomic model was performed in five steps by running molecular-dynamics simulations to form the glass model used for the rest of the simulations. The first step of the melt-and-quench process can be seen in figure 10. The “velocity” command tells the simulation to set the velocities of all the atoms in the simulation box to a specified value that corresponds to a certain temperature [18]. The “all create 3000.0 4545 rot yes dist gaussian” sets the velocity of all the particles in the simulation box to correspond to target temperature of 3000 K with zeroed angular momentum. The “run”

command tells the simulation to perform molecular dynamics for a specified number of timesteps, and the “0” refers to zero timestep, that means to only compute the thermodynamics without running the simulation for any further steps, and with the “all scale 3000.0” command all the velocities of the atoms are scaled to the target temperature of 3000 K. [18]

In the input file, any “fix” command specifies certain attributes related to the simulation. The “1 all nvt temp 3000.0 3000.0 0.04” refers to applying to all of the particles in the simulation box a Nosé-Hoover thermostat to keep the particle number, volume and temperature constant with 3000 K being the initial and final (target) temperature and 0.04 the coupling time of the thermostat in ps. The “timestep” command tells the simulation the value of the timestep for the molecular-dynamics simulation, and the “0.001” corresponds to a timestep of 0.001 ps (1 fs). The “dump” command tells the simulation to store certain quantities related to the simulation at a chosen file (specified by the user) between certain amount of timesteps, and the “1 all xyz 100 mix-GST. xyz” refers to saving all the positions of the atoms in *x*, *y, z* format between every 100 timesteps to the file named “mix-GST.xyz”. The “dump_modify” command tells the simulation to modify the previous “dump” command, and store the atomic coordinates using as relevant elements Ge, Sb and Te for the atomic species. In the following “run” command the simulation is set to run for 30000 timesteps, corresponding to a 30 ps molecular-dynamics simulation. The “unfix” command tells the simulation to remove the previously defined “fix” command, and the “1” refers to the ID of the relevant melt-and-quench step. The “undump” command tells the simulation to turn off the “dump” command that was defined previously, and “1” refers to the ID of the relevant melt-and-quench step. [18] It is noted that, in all of the different steps of the process the number after “unfix” and “undump” indicates the ID number of the relevant step.

```
#Molecular Dynamics melt-and-quench
################################################################################

restart         5000 rst.*.gst

velocity        all create 3000.0 4545 rot yes dist gaussian
run             0
velocity        all scale 3000.0

fix             1 all nvt temp 3000.0 3000.0 0.04
timestep        0.001
dump            1 all xyz 100 mix-GST.xyz
dump_modify     1 element Ge Sb Te
run             30000
unfix           1
undump          1
```

Figure 10 First step of the melt-and-quench process.

In the following steps of the melt-and-quench process, only the main and important commands are mentioned and their changes from the previous steps, in order to keep the descriptions concise. This is due to the fact that all of the steps include similar commands with only different values.

The second step of the melt-and-quench process can be seen in figure 11. With the “all create 1200.0 1212 rot yes dist gaussian” the velocity of all the atoms in the simulation box was set to correspond to target temperature of 1200 K, obtaining a liquid structure at this temperature. Then with the “2 all nvt temp 1200.0 1200.0 0.04”, the particle number, volume and temperature was kept constant using a Nosé-Hoover thermostat, with the initial and final temperature of 1200 K, and a coupling time of the thermostat of 0.04 ps. With the “0.001” the timestep was set to 0.001 ps (1 fs). With the “2 all xyz 50 1-GST.xyz” all the positions of the atoms in *x*, *y, z* format, in between every 50 timesteps, were stored to the file named “1-GST.xyz”. With the “30000” the simulation was set to run for 30000 timesteps, corresponding to a 30 ps molecular-dynamics simulation.

```
velocity        all create 1200.0 1212 rot yes dist gaussian
run             0
velocity        all scale 3000.0

fix             2 all nvt temp 1200.0 1200.0 0.04
timestep        0.001
dump            2 all xyz 50 l-GST.xyz
dump_modify     2 element Ge Sb Te
run             30000
unfix           2
undump          2
```

Figure 11 Second step of the melt-and-quench process.

The third step of the melt-and-quench process can be seen in figure 12. With the "3 all nvt temp 1200.0 300.0 0.04" the number of particles, volume and energy was kept constant using a Nosé-Hoover thermostat, with the initial temperature of 1200 K and the final temperature of 300 K and a coupling time of 0.04 ps to let the liquid structure (1200 K) to cool down to room temperature (300 K), and generate the glass structure. With the "0.001" the timestep was set to 0.001 ps (1 fs). With the "3 all xyz 100 a-GST.xyz" all the positions of the atoms in *x*, *y, z* format, in between each 100 timesteps, were stored to the file named "a-GST.xyz". With the "60000" the simulation was set to run for 60000 timesteps, corresponding to a 60 ps molecular-dynamics simulation.

```
fix             3 all nvt temp 1200.0 300.0 0.04 #cool down to 300K with 15K/ps
timestep        0.001
dump            3 all xyz 100 a-GST.xyz
dump_modify     3 element Ge Sb Te
run             60000
unfix           3
undump          3
```

Figure 12 Third step of the melt-and-quench process.

The fourth step of the melt-and-quench process can be seen in figure 13. With the "4 all nvt temp 300.0 300.0 0.04" the number of particles, volume and temperature was kept constant using a Nosé-Hoover thermostat, with the initial and final temperature of 300 K, and a coupling time of the thermostat of 0.04 ps. With the "0.001" the timestep was set to 0.001 ps (1 fs). With the "4 all xyz 50 glass_NVT_GST.xyz" all the positions of the atoms in *x*, *y, z* format, in between every 50 timesteps, were saved to the file named "glass_NVT_GST.xyz". With the

"30000" the simulation was set to run for 30000 timesteps, corresponding to a 30 ps molecular-dynamics simulation.

```
fix             4 all nvt temp 300.0 300.0 0.04
timestep        0.001
dump            4 all xyz 50 glass_NVT_GST.xyz
dump_modify     4 element Ge Sb Te
run             30000
unfix           4
undump          4
```

Figure 13 Fourth step of the melt-and-quench process.

The fifth and last step of the melt-and-quench process can be seen in figure 14. With the "5 all nve" the number of particles, volume and energy was kept constant. With the "0.001" the timestep was set to 0.001 ps (1 fs). With the "5 all xyz 5 glass_NVE_GST.xyz" all the positions of the atoms in *x*, *y, z* format, in between every 5 timesteps, were stored to the file named "glass_NVE_GST.xyz". With the "10000" the simulation was set to run for 10000 timesteps, corresponding to a 10 ps molecular-dynamics simulation. This step is also called the collection data run, gathering information for further post-processing structural analysis.

```
fix             5 all nve
timestep        0.001
dump            5 all xyz 5 glass_NVE_GST.xyz
dump_modify     5 element Ge Sb Te
run             10000
unfix           5
undump          5
```

Figure 14 Fifth step of the melt-and-quench process.

4.2 Model structure file

The model structure file is the second file needed for the molecular-dynamics simulation using LAMMPS. Figure 15 shows the main part of the model structure file used in the simulations. In the model structure file, the atomic structure is defined by first writing the total number of

the atoms in the model in the first line, with "900" referring to nine hundred atoms in total. Then, the size of the simulation cell in *x*, *y* and *z* directions, as well as the angles of its edges are defined. The value of "30.722" refers to the length of each edge of the simulation box in Å, and the value of "90.0" refers to the angle of each edge in degrees, corresponding to a cubic simulation box. Lastly, separate lines containing the element (atomic species) and its position (*x*, *y,* z coordinates) are included in the model structure file. For example, starting from the third line "Ge" refers to the germanium element, "26.1195" is the *x*- coordinate, "25.6079" is the *y*-coordinate and "21.3254" is the *z*-coordinate.

```
900
30.722 30.722 30.722 90.0 90.0 90.0
Ge 26.1195 25.6079 21.3254
Sb 0.816152 10.9923 28.8739
Ge 22.9017 24.8329 9.41031
Sb 21.3243 17.9698 28.0384
Sb 29.3441 27.3467 23.3879
Sb 5.17497 12.2028 4.6853
Ge 8.86225 23.9619 12.4547
Sb 23.0201 11.4338 19.2344
Ge 0.449627 4.16103 30.315
Sb 10.9576 9.88685 20.5423
```

Figure 15 Model structure file showing the total number of atoms in the model, the dimensions of the simulation cell in x, y and z directions and the angles of the simulation cell edges, with the x, y and z coordinates of the first 10 atoms shown as representative example.

4.3 CSC High Performance Computing Facilities

CSC, IT Center for Science is a Finnish based scientific computing and data management center. CSC is focusing on helping and educating researchers, as well as learners and companies about the research computing world. The expertise of CSC covers three different fields: scientific computing, data management, and digitalization of research and education. The vision is co-operative, building world-class environments for innovation, learning and research. The customer segment includes several different parties: universities, research institutions and infrastructures, public administration, companies, archives, libraries and museums. High performance computing is a research instrument, which is necessary for data-intensive and computing research. The high performance computing facilities of CSC currently include three different supercomputers: Mahti, Puhti and LUMI, which are all located in Kajaani, Finland. Mahti is a robust supercomputer, based on BullSequana XH2000 system, that has been

launched in the national computing environment of CSC in 2020. The usage of Mahti is intended for medium to large scale simulations, which require high computing performance and a fast interconnection network. Mahti-AI, which is the GPU-accelerated partition of Mahti is especially intended for big data processing and AI-related research. The CPU partition capacity of the machine can reach up to 7.5 petaflops, also called as 7.5 quadrillion operations per second. Mahti-AI, the GPU side, reaches up to 2.0 petaflops per second. Puhti is a supercomputer, based on X400 system, for general purpose computing tasks. It has been launched in 2019 and it is suitable for a wide range of different usages, from interactive data analysis all the way to medium scale simulations. Puhti has a wide range of different scientific software available for scientists to use. Puhti has also its partition for AI applications, called Puhti-AI. The CPU partition of Puhti reaches up to 1.8 petaflops per second and the GPU-accelerated AI side reaches up to 2.7 petaflops per second. LUMI is a unique European supercomputer that is one of the most powerful in the whole world, based on HPE Cray EX architecture. LUMI provides a state-of-the-art computing performance for a wide range of usage and problems, such as accurate climate modeling, for instance. LUMI reaches up to 380 petaflops per second, corresponding to 380 quadrillion operations each second. This computing power equals to the performance of 1.5 million laptop computers. LUMI´s largest partition is in the GPU-accelerated LUMI-G, but has also smaller partitions in CPU based, namely, LUMI-C and LUMI-D. [32]

4.4 Study workflow

The workflow of the study was as follows. The first step was to build up the needed input file, as well as the model structure file. Then both files were uploaded into the CSC high performance computing facilities server, and simulations were set to be started. After all of the simulations were completed, all the calculation data were gathered from the server to be analyzed. In the analysis, all the simulation data, regarding temperature profiles, local atomic structure and atomic mobility were extracted for further processing. Then, by using the extracted simulation data, multiple properties were calculated and relevant figures and tables were generated.

5 Results

From the simulations, information about local atomic structure and atomic mobility were gathered in three different temperature regions. Also, data about the temperature fluctuations during the simulations were analyzed accordingly.

5.1 Temperature profiles

From the simulation log files, information about the temperature of the model system was able to be extracted, that worked as quality check proof that the molecular-dynamics simulation was able to reproduce the target temperature at a satisfactory level. The figure 16 presents the temperature fluctuations from the simulations in the overheated amorphous temperature region. One can see that the temperature fluctuates approximately ± 20 K around the target temperature of 420 K, providing an indication of the validity of the relevant simulation.

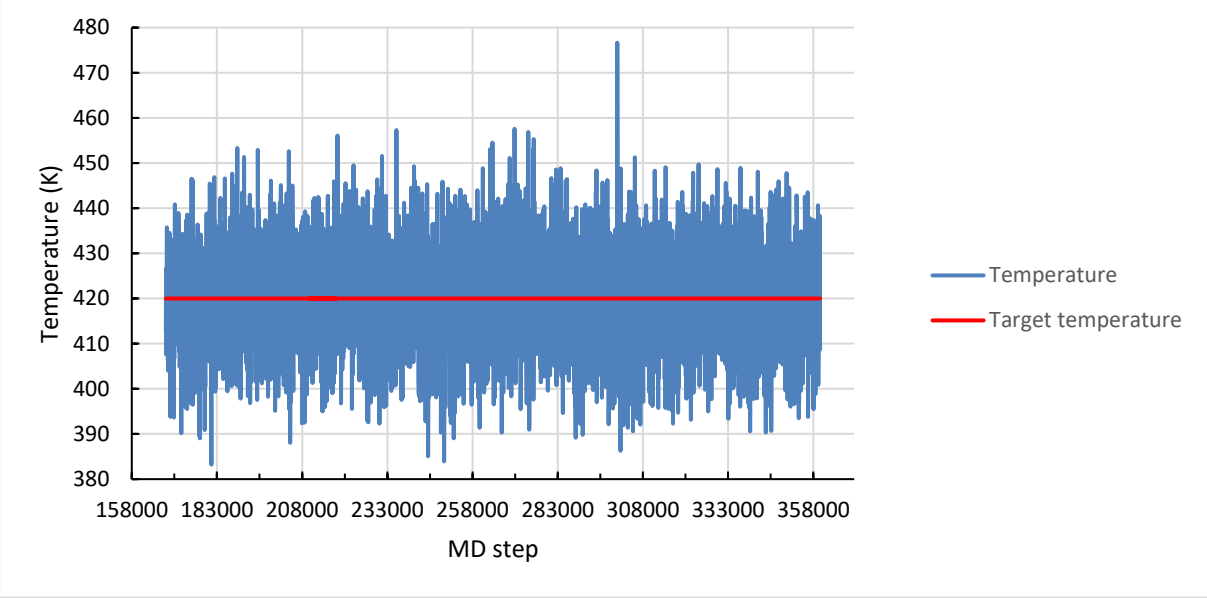

Figure 16 Temperature fluctuations around the target temperature of 420 K during the simulations in the overheated amorphous temperature region.

The figure 17 presents the temperature fluctuations of the model system in the supercooled liquid temperature region. One can see that the temperature fluctuates approximately ± 37.5 K around the target temperature of 825 K, providing an indication of the validity of the simulation at the supercooled liquid temperature.

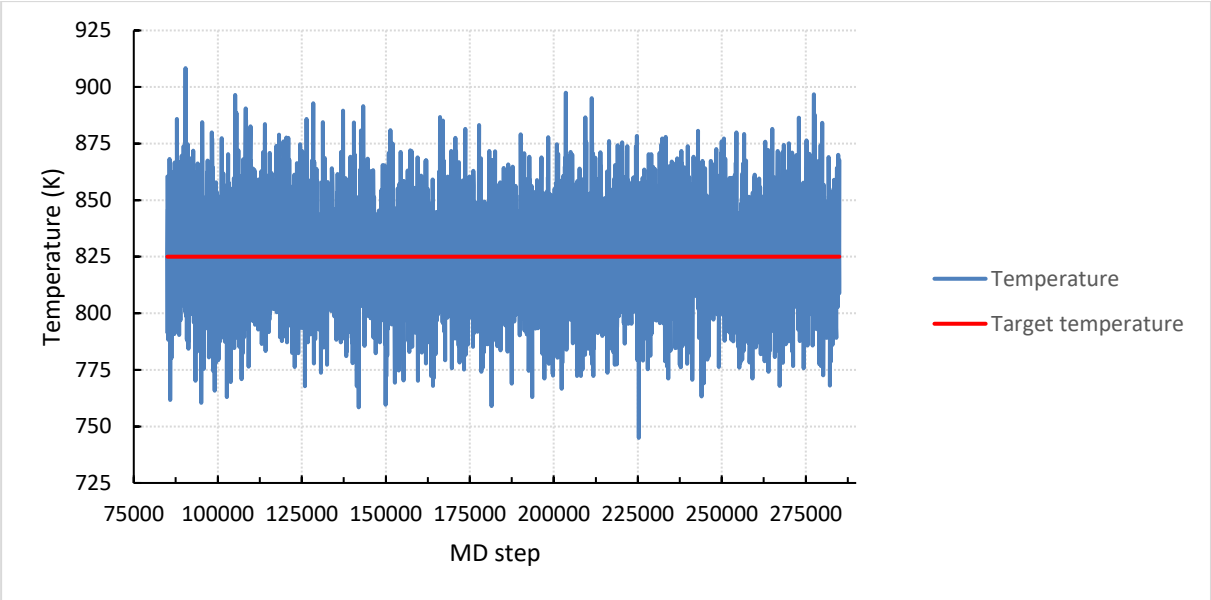

Figure 17 Temperature fluctuations around the target temperature of 825 K during the simulations in the supercooled liquid temperature region.

The figure 18 presents the temperature fluctuations of the model system in the liquid temperature region. One can see that the temperature fluctuates approximately ± 60 K around the target temperature of 1200 K, providing an indication of the validity of the simulation at the liquid temperature.

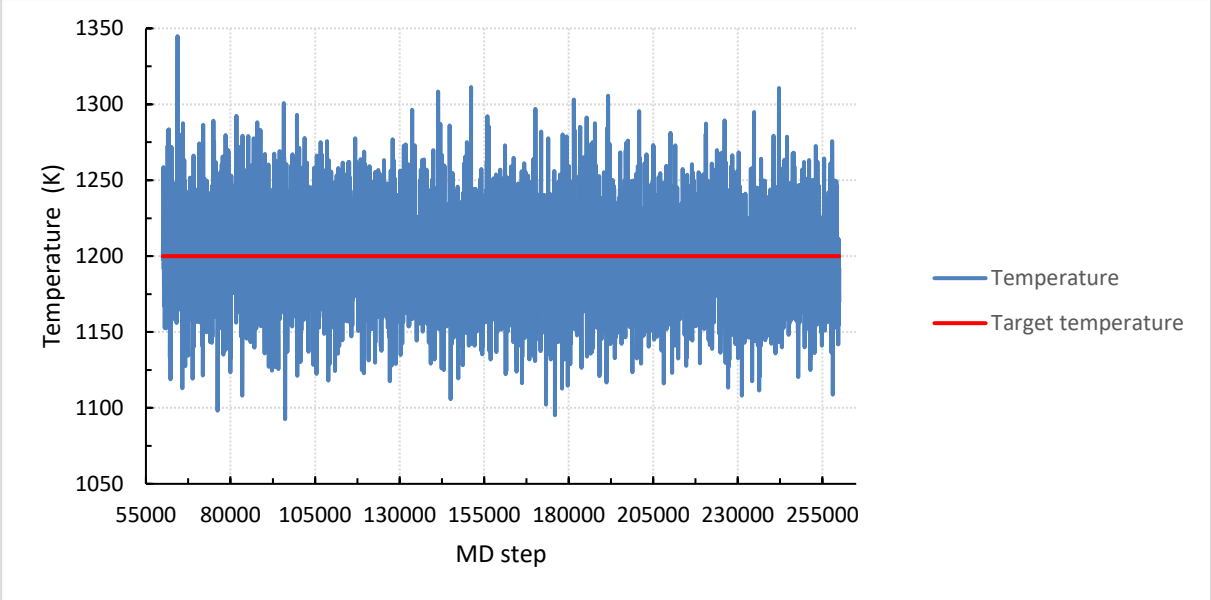

Figure 18 Temperature fluctuations around the target temperature of 1200 K during the simulations in the liquid temperature region.

The image in figure 19 shows the 900 atoms atomistic model that was used in the simulations performed in this thesis. Blue colored spheres represent Ge atoms, red ones Sb atoms, and yellow ones Te atoms. The image was taken as a snapshot by using the visualization software VMD [33].

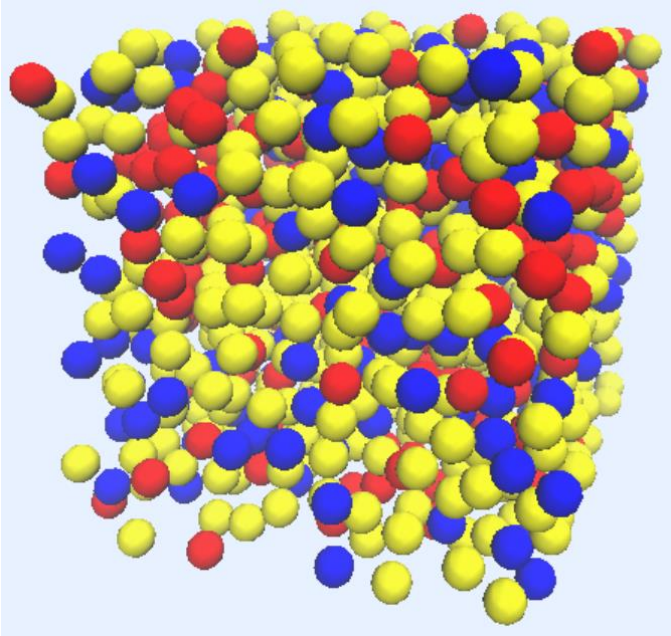

Figure 19 900-atom $Ge_2Sb_2Te_5$ atomistic model used in the simulations.

5.2 Local atomic structure

The analysis associated with the local atomic structure of the simulated models includes information about radial distribution function, coordination numbers and bond angle distributions in the three different temperature regions studied in this thesis.

5.2.1 Radial Distribution Functions

The figure 20 presents the radial distribution function as a function of distance in Å at the temperature of 420 K for five different atom pairs, Ge-Ge, Ge-Te, Sb-Sb, Sb-Te and Te-Te. The figure provides physical information about the interatomic distance between the certain atoms, i.e., their bond length. From the figure, one can see that the distance with the highest probability of finding germanium-germanium pair is at 3.95 Å, for germanium-tellurium pair at 2.82 Å, for antimony-antimony pair at 4.21 Å, for antimony-tellurium pair at 3.00 Å, and for tellurium-tellurium pair at 4.13 Å. Also, one can see that the longer the distance goes, over 8

Å, the probability of finding each atom pair is vanishing, because these distances correspond to regions beyond the first-coordination shell of a given atom.

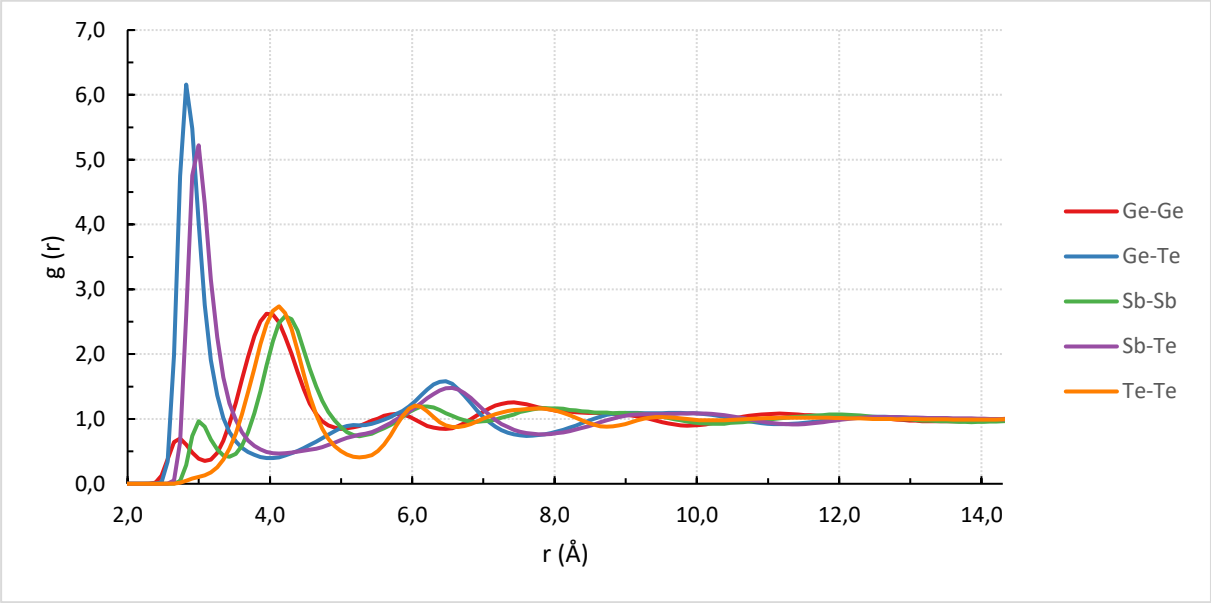

Figure 20 Radial distribution functions of selected atom pairs in the simulated $Ge_2Sb_2Te_5$ structure at the temperature of 420 K.

The figure 21 presents the radial distribution functions at the temperature of 825 K for five different atom pairs, Ge-Ge, Ge-Te, Sb-Sb, Sb-Te and Te-Te. From the figure, one can see that the distance with the highest probability of finding germanium-germanium pair is 3.95 Å, for germanium-tellurium pair at 2.82 Å, for antimony-antimony pair at 3.00 Å, for antimony-tellurium pair at 3.00 Å, and for tellurium-tellurium pair at 4.04 Å.

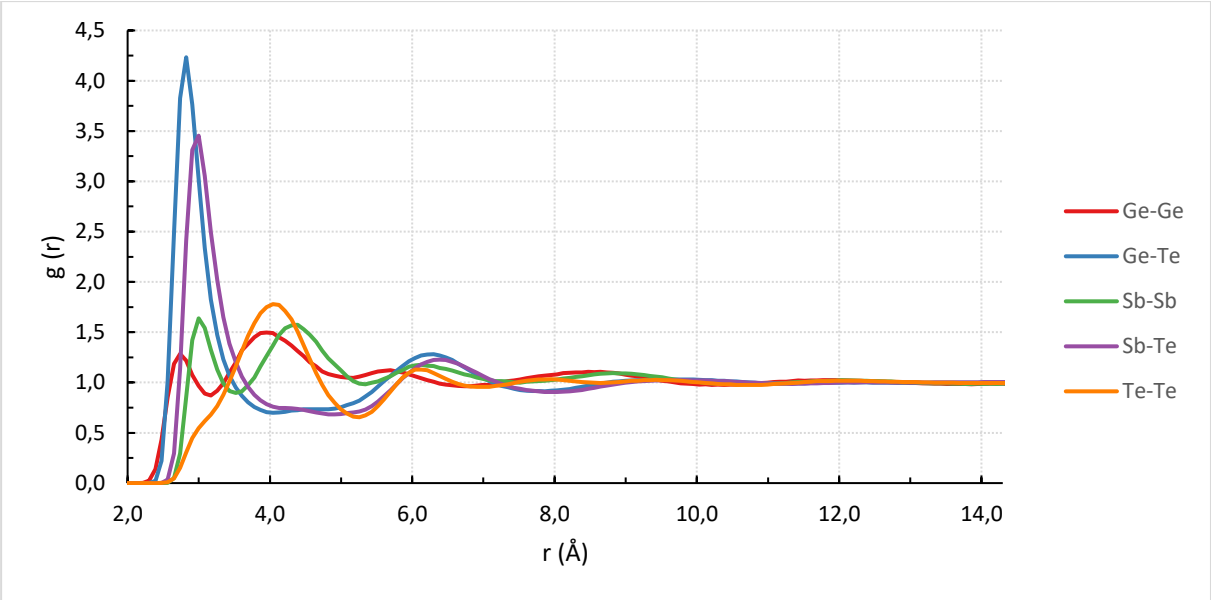

Figure 21 Radial distribution functions of selected atom pairs in the simulated $Ge_2Sb_2Te_5$ structure at the temperature of 825 K.

The figure 22 presents the radial distribution functions at the temperature of 1200 K for five different atom pairs, Ge-Ge, Ge-Te, Sb-Sb, Sb-Te and Te-Te. From the figure, one can see that the distance with the highest probability of finding germanium-germanium pair is 2.74 Å, for germanium-tellurium pair at 2.82 Å, for antimony-antimony pair at 3.00 Å, for antimony-tellurium pair at 3.00 Å, and for tellurium-tellurium pair at 3.87 Å.

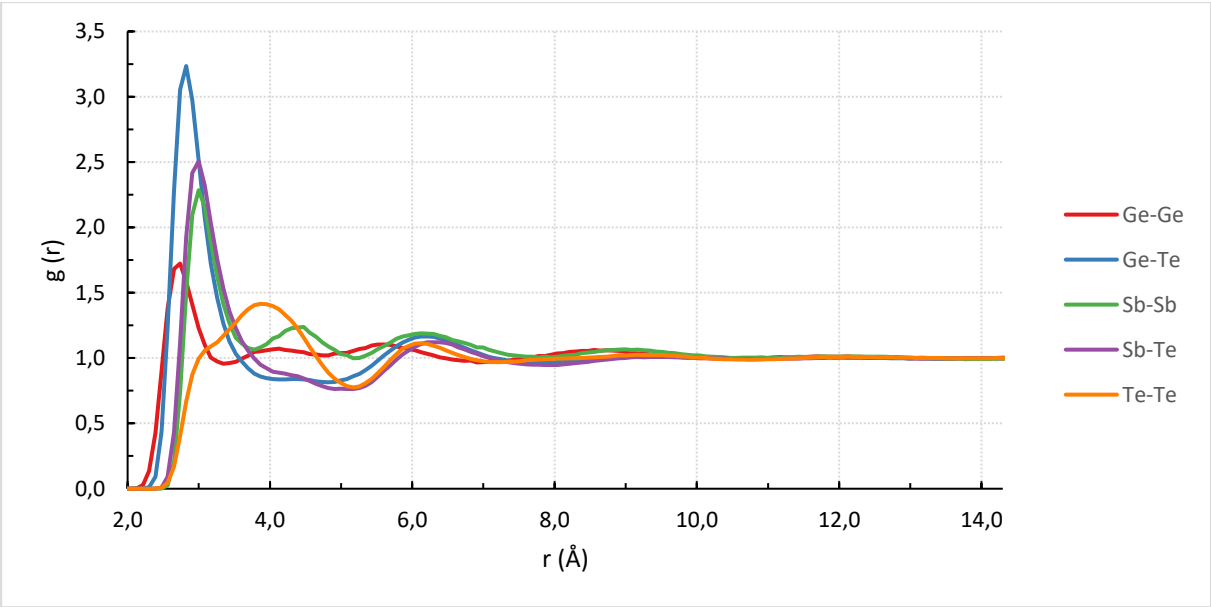

Figure 22 Radial distribution functions of selected atom pairs in the simulated $Ge_2Sb_2Te_5$ structure at the temperature of 1200 K.

The figure 23 presents the radial distribution functions of germanium-germanium pair in three different temperatures of 420 K, 825 K, and 1200 K. One can see, observing the behaviour of the first coordination shell, that the higher the temperature is the higher the probability of finding germanium-germanium pair or, in other words, the higher the amount of Ge-Ge formations within the simulated structure.

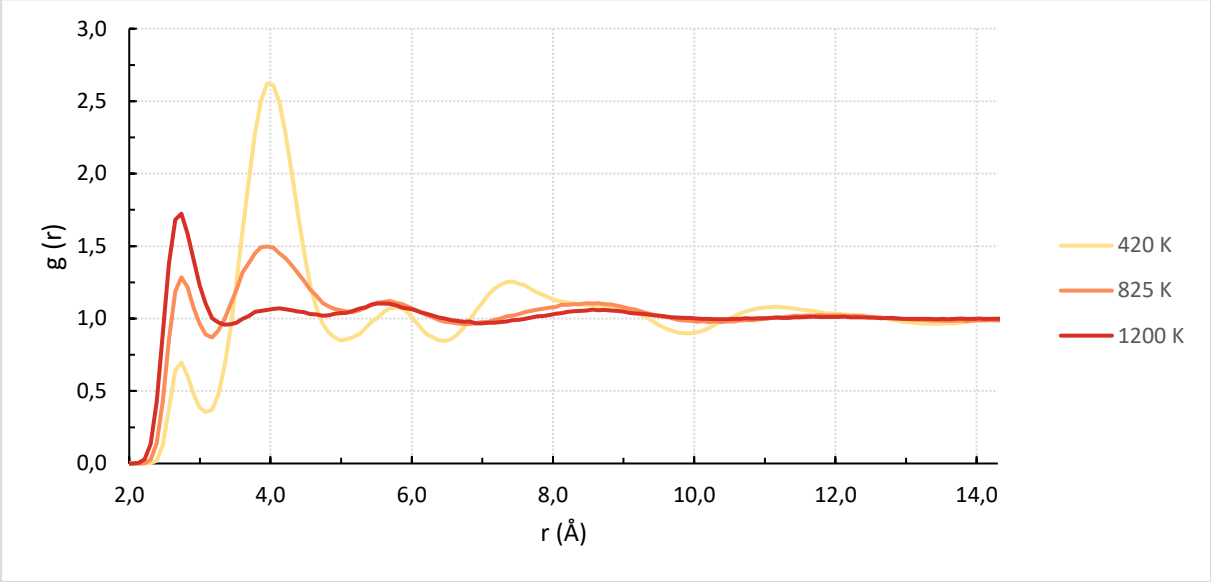

Figure 23 Radial distribution functions of germanium-germanium pair in three different temperatures.

The figure 24 presents the radial distribution functions of germanium-tellurium pair in three different temperatures of 420 K, 825 K, and 1200 K. One can see, observing the behaviour of the first coordination shell, that the higher the temperature is the lower the probability of finding germanium-tellurium pair or, in other words, the lower the amount of Ge-Te formations inside the $Ge_2Sb_2Te_5$ model.

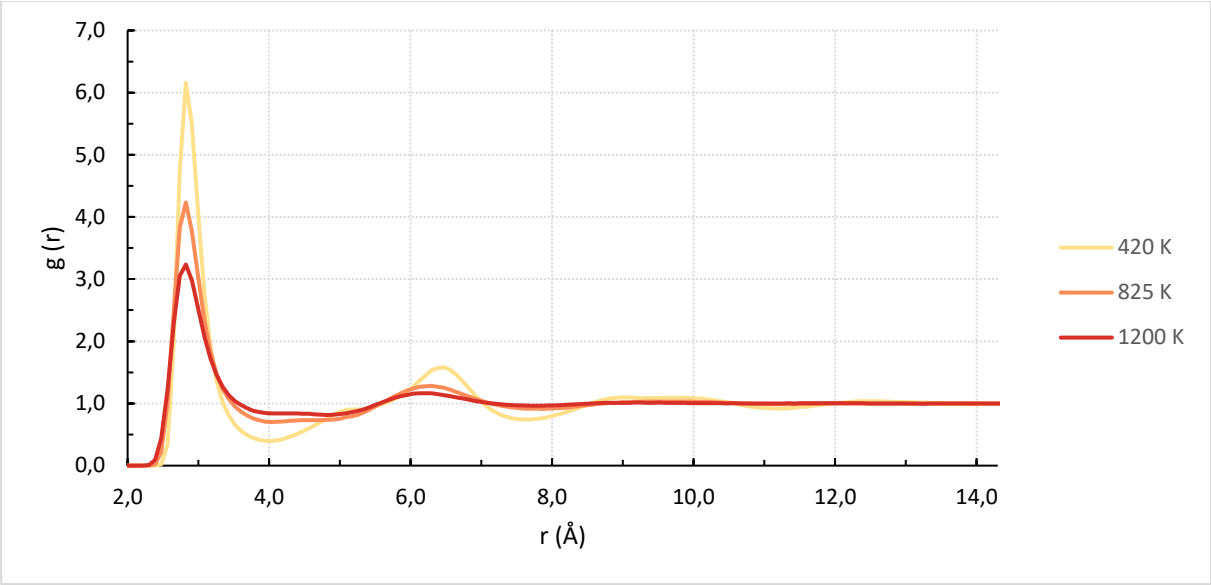

Figure 24 Radial distribution functions of germanium-tellurium pair in three different temperatures.

The figure 25 presents the radial distribution functions of antimony-tellurium pair in three different temperatures of 420 K, 825 K, and 1200 K. One can see, observing the behaviour of the first coordination shell that the higher the temperature is the lower the probability of finding antimony-tellurium pair or, in other words, the lower the amount of Sb-Te formations in the simulated structure.

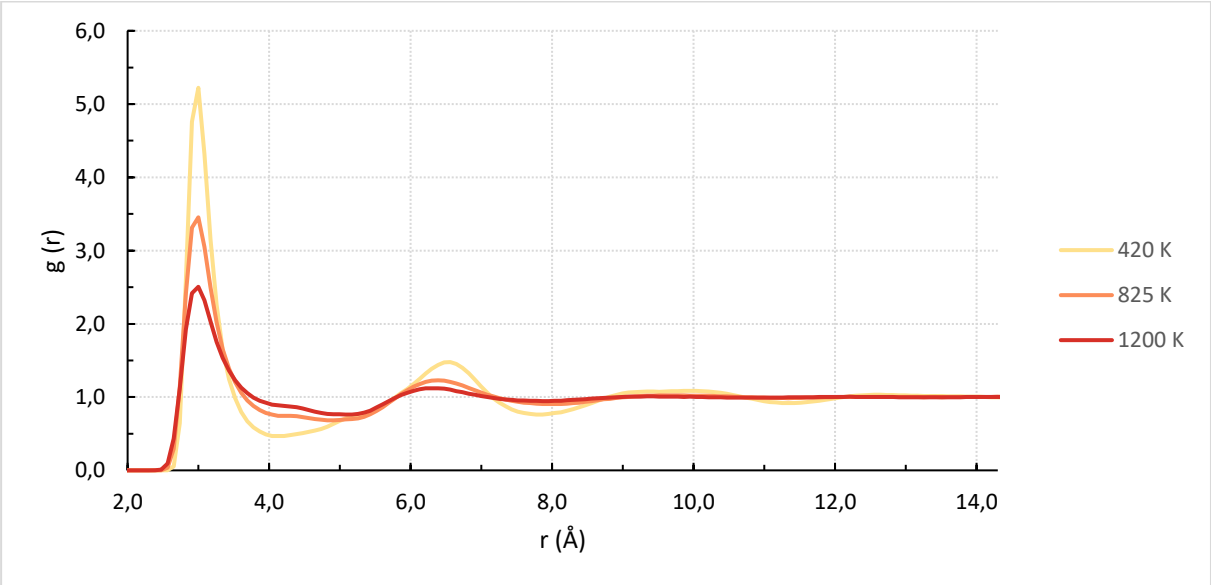

Figure 25 Radial distribution functions of antimony-tellurium pair in three different temperatures.

The figure 26 presents the radial distribution functions of tellurium-tellurium pair in three different temperatures of 420 K, 825 K, and 1200 K. One can see, observing the behaviour of

the first coordination shell that the higher the temperature is the lower the probability of finding tellurium-tellurium pair or, in other words, the lower the amount of Te-Te formations in the $Ge_2Sb_2Te_5$ model.

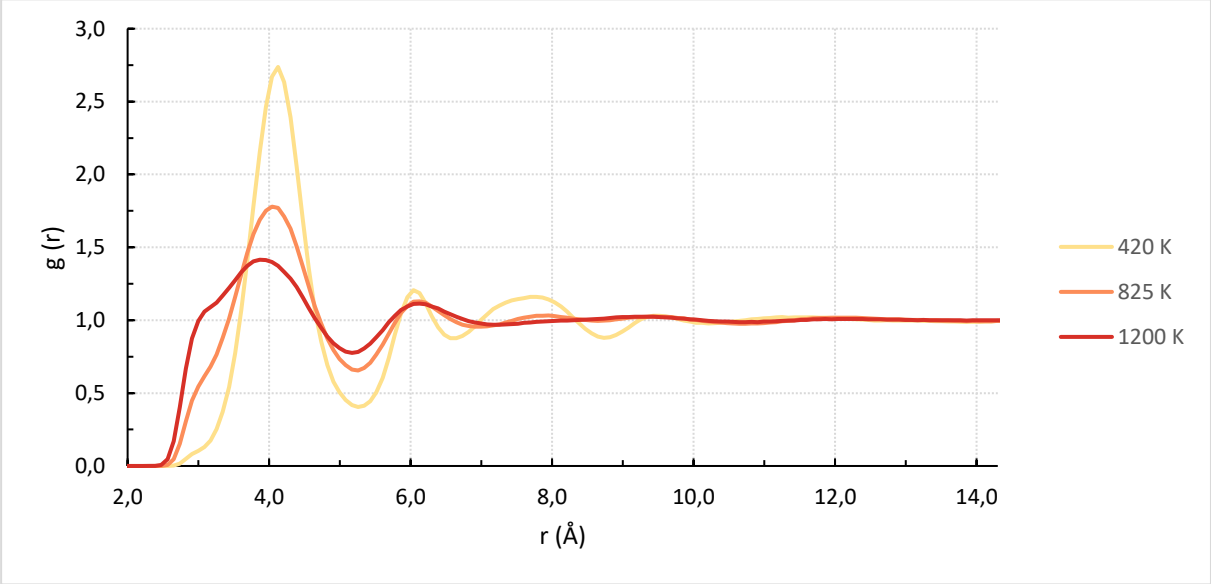

Figure 26 Radial distribution functions of tellurium-tellurium pair in three different temperatures.

The figure 27 presents the radial distribution functions of antimony-antimony pair in three different temperatures of 420 K, 825 K, and 1200 K. One can see, observing the behaviour of the first coordination shell that the higher the temperature is the higher the probability of finding antimony-antimony pair or, in other words, the higher the amount of Sb-Sb formations in the simulated structure.

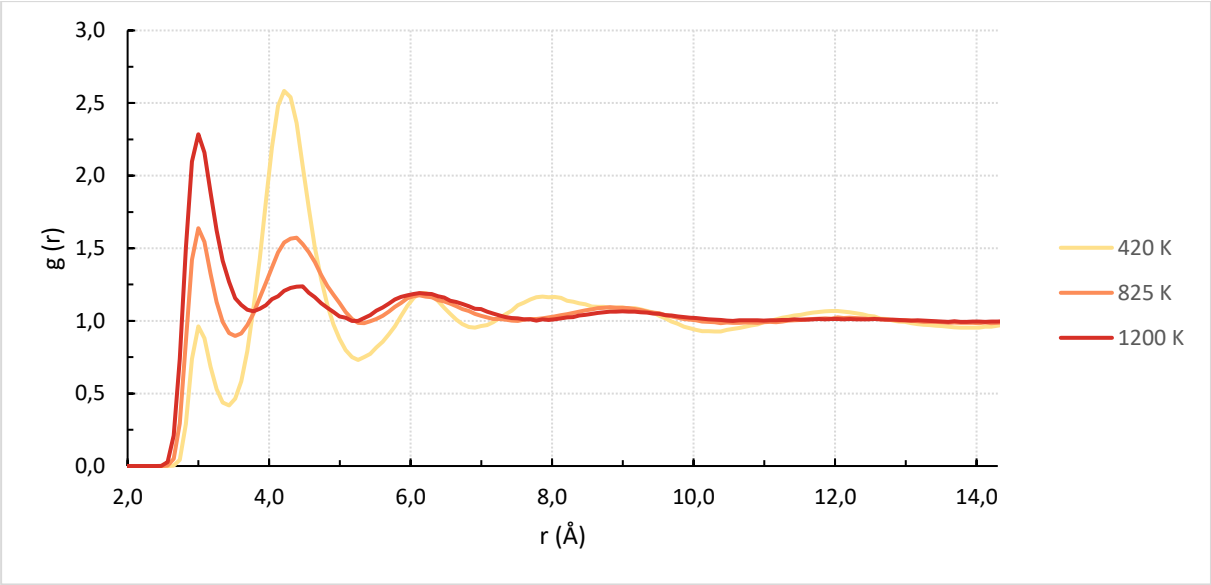

Figure 27 Radial distribution functions of antimony-antimony pair in three different temperatures.

From the table 2, one can see the bond lengths of the element pairs in Å in the three different temperatures studied in this thesis. The bond length of germanium-germanium pair is 3.95 Å at 420 K, as well as at 825 K. At the temperature of 1200 K the bond length of germanium-germanium pair gets shorter to 2.74 Å. The bond length of germanium-tellurium pair is 2.82 Å in all three temperatures (420 K, 825 K, and 1200 K). Similarly, the bond length of antimony-tellurium pair is 3.00 Å in all three temperatures. The bond length of antimony-antimony pair is 4.21 Å at 420 K and gets shorter to 3.00 Å at the higher temperatures of 825 K and 1200 K. The bond length of tellurium-tellurium pair becomes shorter the higher the temperature gets, being 4.13 Å at 420 K, 4.04 Å at 825 K, and 3.87 at 1200 K.

Table 2 Calculated bond lengths of all the element pairs at the three different temperatures.

	Bond length (Å)		
Atom pair	420 K	825 K	1200 K
Ge-Ge	3.95	3.95	2.74
Ge-Te	2.82	2.82	2.82
Sb-Te	3.00	3.00	3.00
Sb-Sb	4.21	3.00	3.00
Te-Te	4.13	4.04	3.87

5.2.2 Coordination Numbers

Table 3 presents the coordination numbers of each element, germanium, antimony and tellurium, in the three different temperatures. From the table, one can see that the coordination number gets lower the higher the temperature gets. The largest coordination number deviation is when the temperature rises from 420 K to 825 K. From 825 K to 1200 K the coordination number does not increase any further, staying at similar values for all the atomic species.

Table 3 Coordination numbers of germanium, antimony and tellurium atoms in the three different temperatures.

	Coordination number		
Atom	420 K	825 K	1200 K
Ge	4.6	4.3	4.3
Sb	3.8	3.4	3.3
Te	3.2	2.9	2.9

The figure 28 presents the coordination numbers distribution of germanium in the three different temperatures of 420 K, 825 K and 1200 K. The percentage of 2-coordinated Ge atoms at the temperatures of 825 K and 1200 K are between 0 % to 5 %, whereas at 420 K such coordination is negligible. The percentage of 3-coordinated Ge atoms at the temperature of 825 K and 1200 K is between 15 % to 20 %, while at 420 K almost 10 % of 3-coordinated Ge atoms are present. The percentage of 4-coordinated Ge atoms at all different temperatures is quite similar, being between 35 % to 40 %. The percentage of 5-coordinated Ge atoms is highest at the temperature of 420 K, being between 35 % to 40 %. At the temperature of 825 K the percentage of 5-coordinated Ge atoms is around 30 %, and at the temperature of 1200 K is just below 30 %. The percentage of 6-coordinated Ge atoms at the temperatures of 825 K and 1200 K is about 10 %, while at the temperature of 420 K this percentage increases to around 15 %. The percentage of 7-coordinated Ge atoms at all three temperatures is very minimal in the simulated structure (being almost negligible). Overall, at all temperatures, there are mostly 4- and 5-coordinated Ge atoms, which correspond to 60 % - 75 % of the local atomic coordination of Ge atoms in the $Ge_2Sb_2Te_5$ simulated structure. The 3- and 6-coordinated Ge atoms correspond to 25 % - 30 % of the environments in the simulated structure.

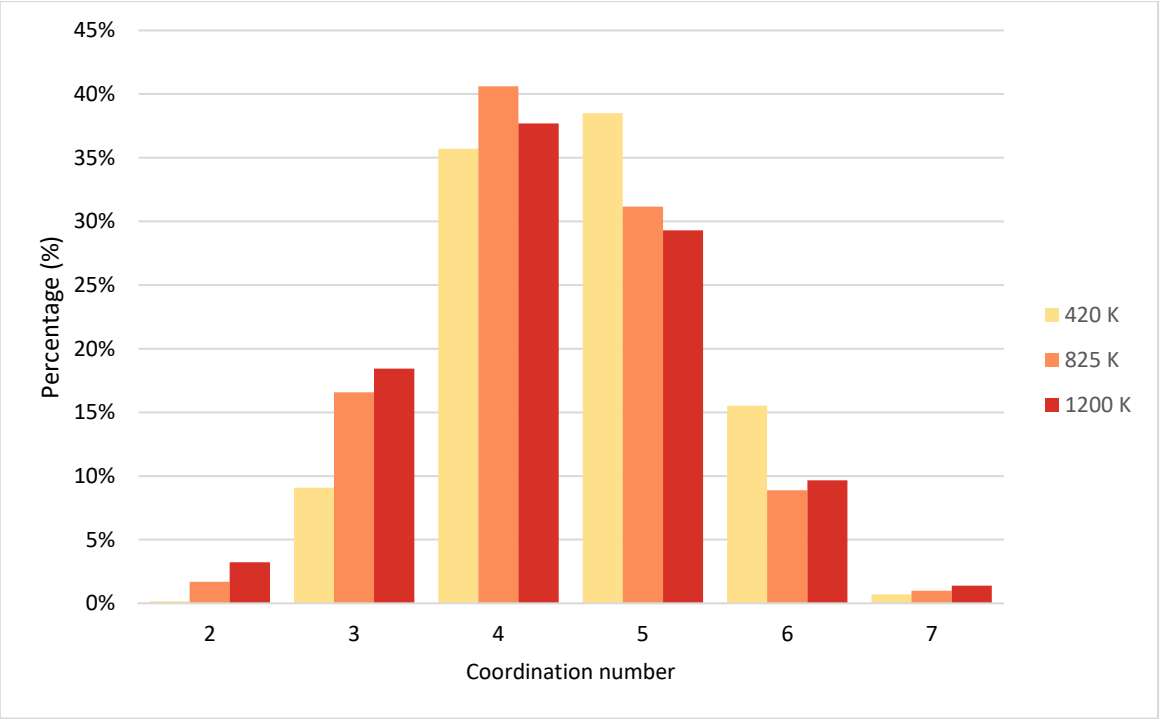

Figure 28 Coordination numbers distribution of germanium atoms in the three different temperatures.

The figure 29 presents the coordination numbers distribution of antimony in the three different temperatures of 420 K, 825 K and 1200 K. The percentage of 2-coordinated Sb atoms at the different temperatures varies, being 5 % at 420 K, almost 15 % at 825 K and slightly over 15 % at 1200 K. The percentage of 3-coordinated Sb atoms at the temperature of 420 K is between 30 % to 35 %, at the temperature of 825 K almost 40 %, and at 1200 K around 35 %. The percentage of 4-coordinated Sb atoms at the temperature of 420 K is slightly over 40 %, at the temperature of 825 K between 30 % to 35 %, and at 1200 K almost 30 %. The percentage of 5-coordinated Sb atoms is the highest at the temperature of 420 K, being between 15 % to 20 %. At both temperatures of 825 K and 1200 K, the percentage of 5-coordinated Sb atoms is around 10 %. The percentage of 6-coordinated Sb atoms at all temperatures is between 0 % to 5 %. Overall, at all temperatures, there are mostly 3- and 4-coordinated Sb atoms, which correspond to around 65 % to 80 % of the local atomic coordination of Sb atoms in the $Ge_2Sb_2Te_5$ simulated structure. The 2- and 5-coordinated Sb atoms correspond to 20 % - 30 % of the environments in the simulated structure.

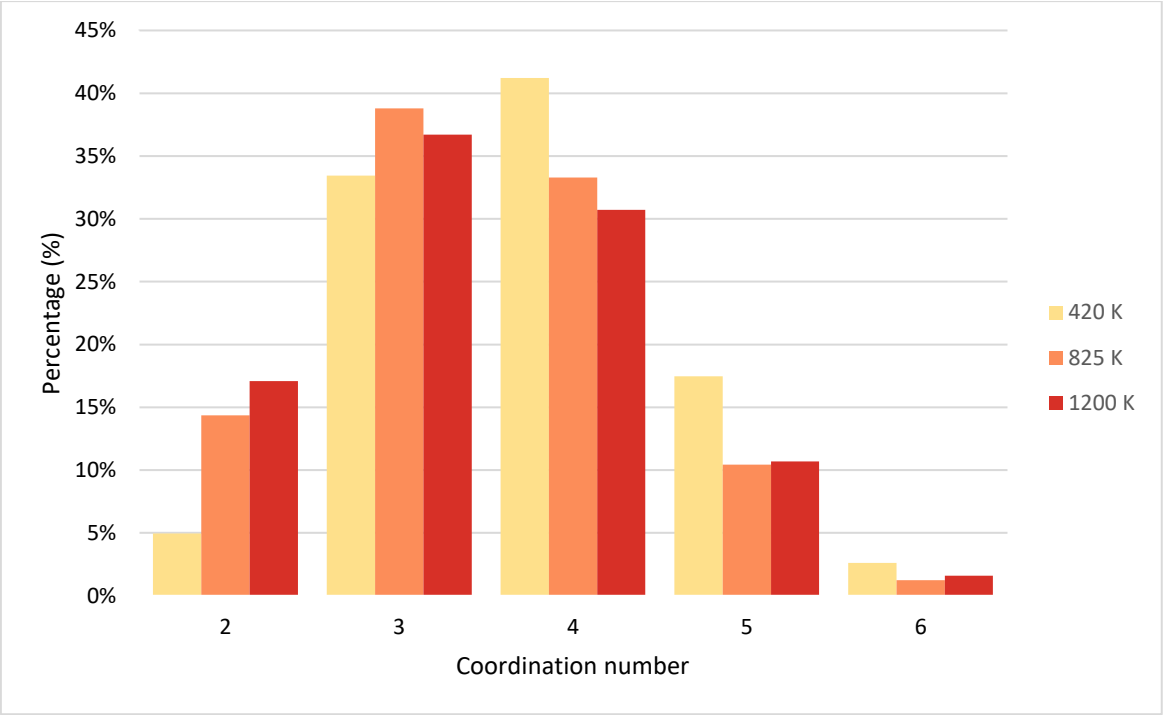

Figure 29 Coordination numbers distribution of antimony atoms in the three different temperatures.

The figure 30 presents the coordination numbers distribution of tellurium in the three different temperatures of 420 K, 825 K and 1200 K. The percentage of 2-coordinated Te atoms at the temperature of 420 K is almost 20 %, while at the temperatures of 825 K and 1200 K is almost 30 %. The percentage of 3-coordinated Te atoms at the temperature of 420 K is 50 %, at the temperature of 825 K is slightly above 40 %, and at 1200 K is 40 %. The percentage of 4-coordinated Te atoms at all temperatures is quite similar, since at the temperature of 420 K is slightly above 20 %, while at the temperatures of 825 K and 1200 K is 20 %. The percentage of 5-coordinated Te atoms is around 5 % at each temperature studied here. Overall, at all temperatures, there are mostly 2-, 3- and 4-coordinated Te atoms, which correspond to around 90 % - 95 % of the local atomic coordination of Te in the $Ge_2Sb_2Te_5$ simulated structure.

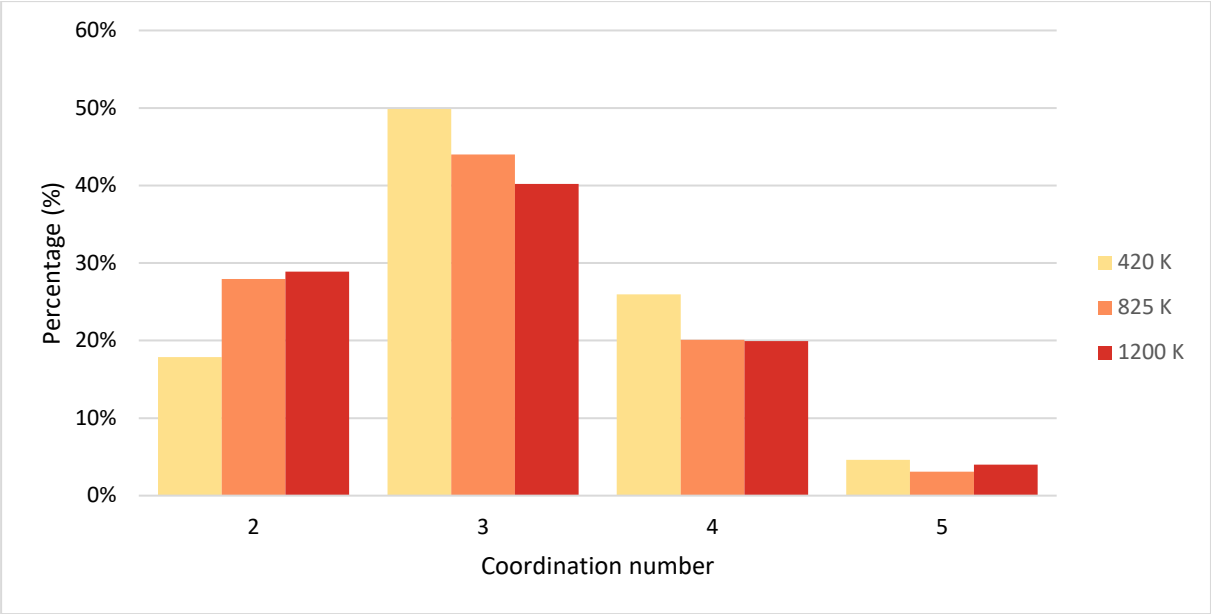

Figure 30 Coordination numbers distribution of tellurium atoms in the three different temperatures.

5.2.3 Bond Angle Distributions

The figure 31 presents the bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium triatomic environments at the temperature of 420 K. From the figure, one can see that both Te-Ge-Te and Te-Sb-Te distributions show two clear angle peaks. The first peak for Te-Ge-Te is at 90.5 degrees, and for Te-Sb-Te is at 87.5 degrees. The second peak in the Te-Ge-Te distribution is at 167.5 degrees, and for the Te-Sb-Te is at 165.5 degrees.

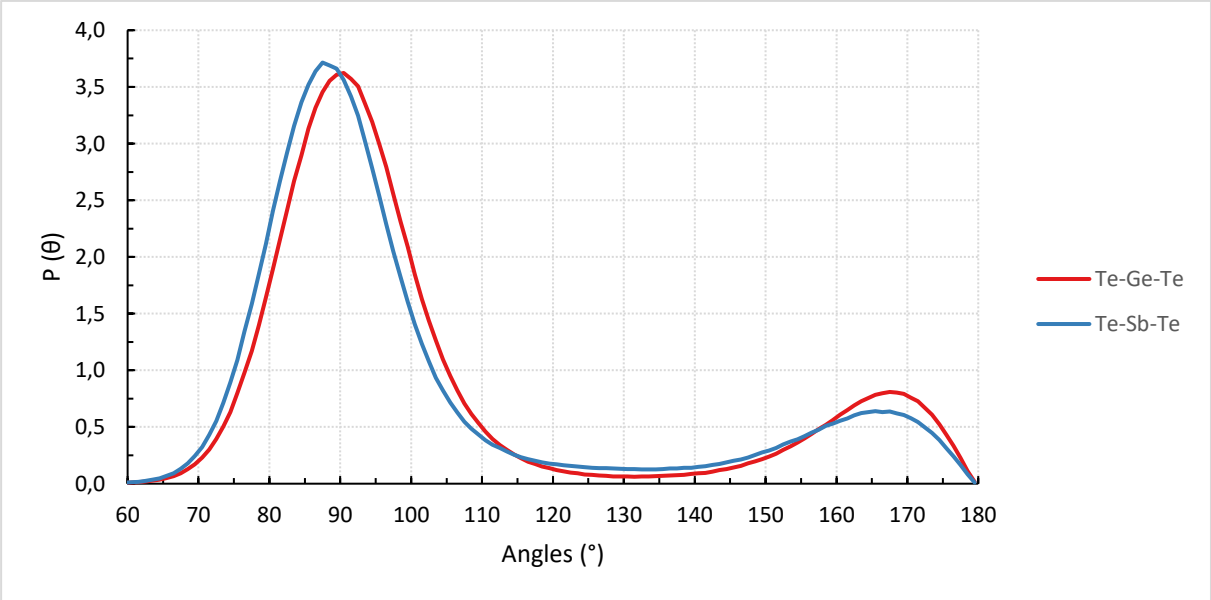

Figure 31 Bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium angles at the temperature of 420 K.

The figure 32 presents the bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium triatomic environments at the temperature of 825 K. From the figure, one can see that both Te-Ge-Te and Te-Sb-Te distributions show again two kinds of angle peaks. The first peak for Te-Ge-Te is at 89.5 degrees and for Te-Sb-Te is at 87.5 degrees. The second peak in the Te-Ge-Te distribution is at 158.5 degrees, and in the Te-Sb-Te is at 146.5 degrees.

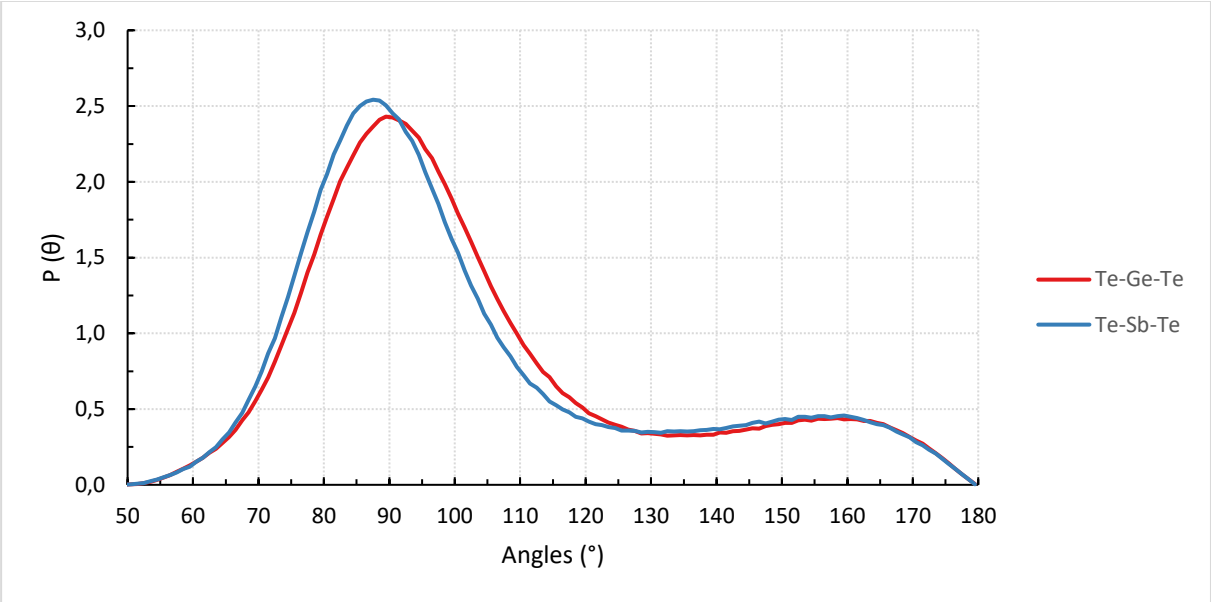

Figure 32 Bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium angles at the temperature of 825 K.

The figure 33 presents the bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium triatomic environments at the temperature of 1200 K. From the figure, one can see that both Te-Ge-Te and Te-Sb-Te distributions form only one clear angle peak. The peak for the Te-Ge-Te angle is at 88.5 degrees, and for the Te-Sb-Te angle is at 87.5 degrees.

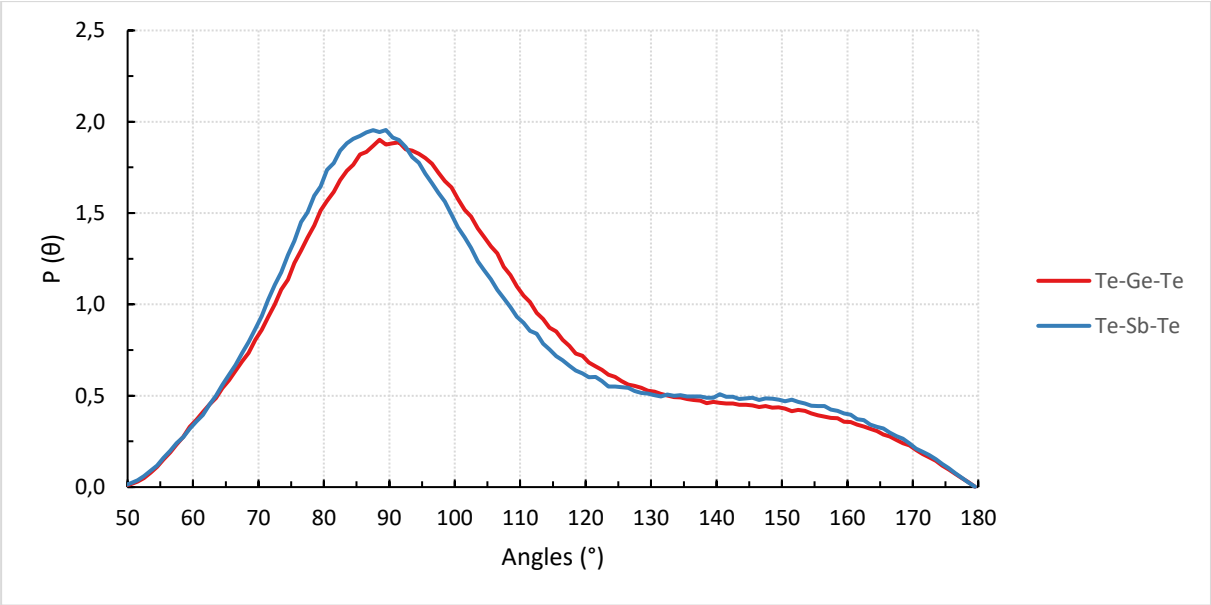

Figure 33 Bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium angles at the temperature of 1200 K.

The figure 34 presents the bond angle distributions of tellurium-germanium-tellurium triatomic environments at the three different temperatures studied here. From the figure, one can see that the first Te-Ge-Te angle peak is about the same at all temperatures, but the lower the temperature gets, the higher the distribution of the angle in the simulated structure becomes. Also, it can be observed that at the region of 150 degrees to 175 degrees, only at the temperature of 420 K a well-defined peak is formed, highlighting that such environments are absent in the two higher temperatures.

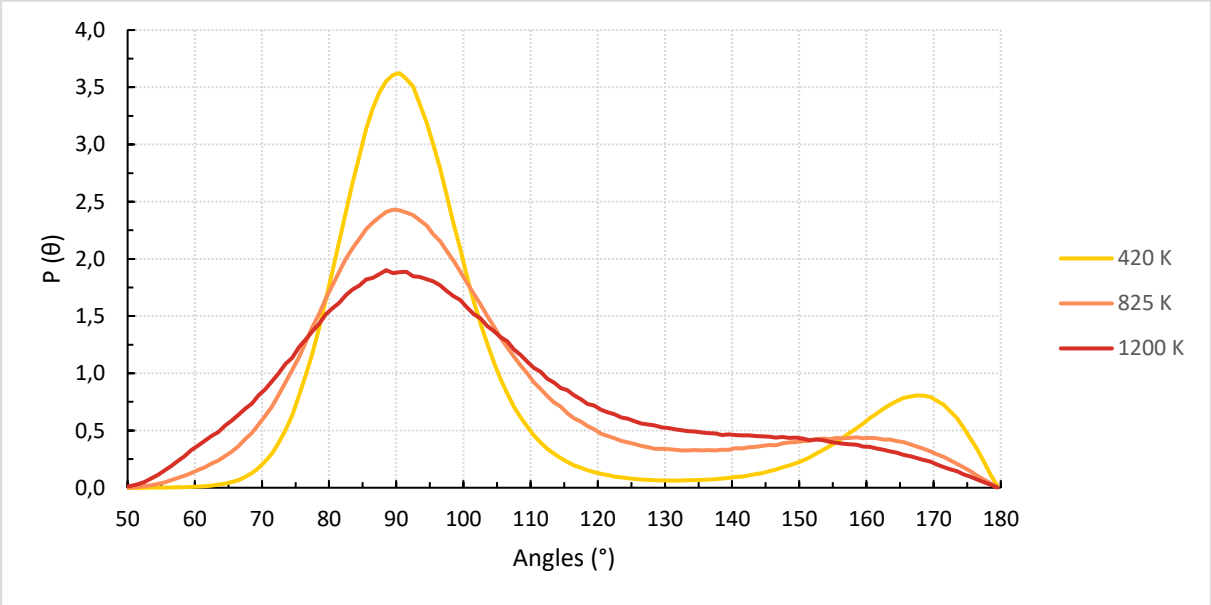

Figure 34 Bond angle distributions of tellurium-germanium-tellurium in the three different temperatures.

The figure 35 presents the bond angle distributions of tellurium-antimony-tellurium triatomic environments at the three different temperatures. From the figure, one can see that the first Te-Sb-Te angle peak is about the same at all temperatures, but the lower the temperature gets, the higher the distribution of the angle becomes in the model. At the angle region of 150 degrees to 175 degrees, one can see that at the temperature of 420 K a well-defined peak is formed, highlighting the existence of such atomic environments at this lower temperature. It can be noted that the Te-Ge-Te and Te-Sb-Te bond angle distributions show a similar behaviour in the simulated $Ge_2Sb_2Te_5$ structure, across the different temperatures studied here.

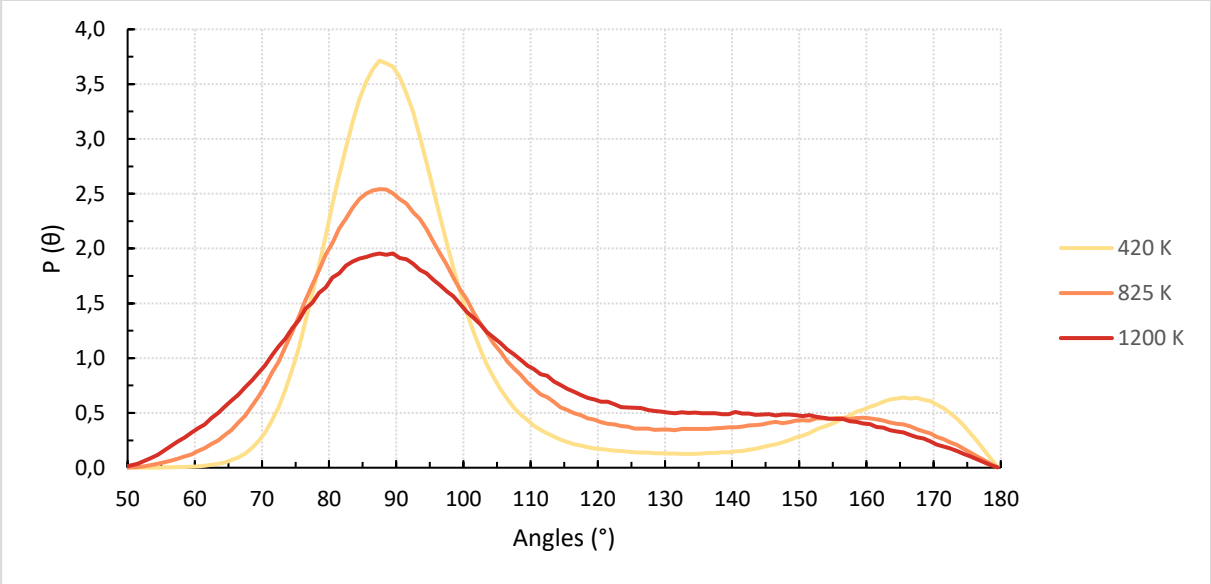

Figure 35 Bond angle distributions of tellurium-antimony-tellurium in the three different temperatures.

Table 4 presents the values of the two identified peaks in the bond angle distributions of Te-Ge-Te and Te-Sb-Te in the three different temperatures studied in this thesis. Noticeable is that the first angle peak, at all three temperatures, is located at similar values for both element combinations. For the second angle peak the corresponding values are more varying for both distributions, showing that the higher the temperature is the lower the angle peak becomes.

Table 4 Two angle peaks in bond angle distributions of tellurium-germanium-tellurium and tellurium-antimony-tellurium in the three different temperatures.

	Two peaks in bond angle distribution					
	1st peak (°)			2nd peak (°)		
Angle	420 K	825 K	1200 K	420 K	825 K	1200 K
Te-Ge-Te	90.5	89.5	88.5	167.5	158.5	139.5
Te-Sb-Te	87.5	87.5	87.5	165.5	146.5	130.5

5.3 Atomic mobility

The analysis associated with the atomic mobility within the simulated $Ge_2Sb_2Te_5$ structure includes information about mean squared displacement and diffusion coefficient of all three elements (Ge, Sb, and Te) in the three different temperature regions studied in this thesis.

5.3.1 Mean Squared Displacement

The figure 36 presents the mean squared displacement of Ge atoms as a function of time, calculated based on the equation (12) from the molecular-dynamics trajectories, in the three different temperatures. From the figure, one can see that at the temperature of 425 K the movement of Ge atoms is really slow, at the temperature of 825 K their movement starts to be moderate, and at the temperature of 1200 K the MSD rapidly increases. The main observation is that the higher the temperature the faster the movement of Ge atoms, which is in accordance with the high mobility of atoms in high (liquid) temperatures.

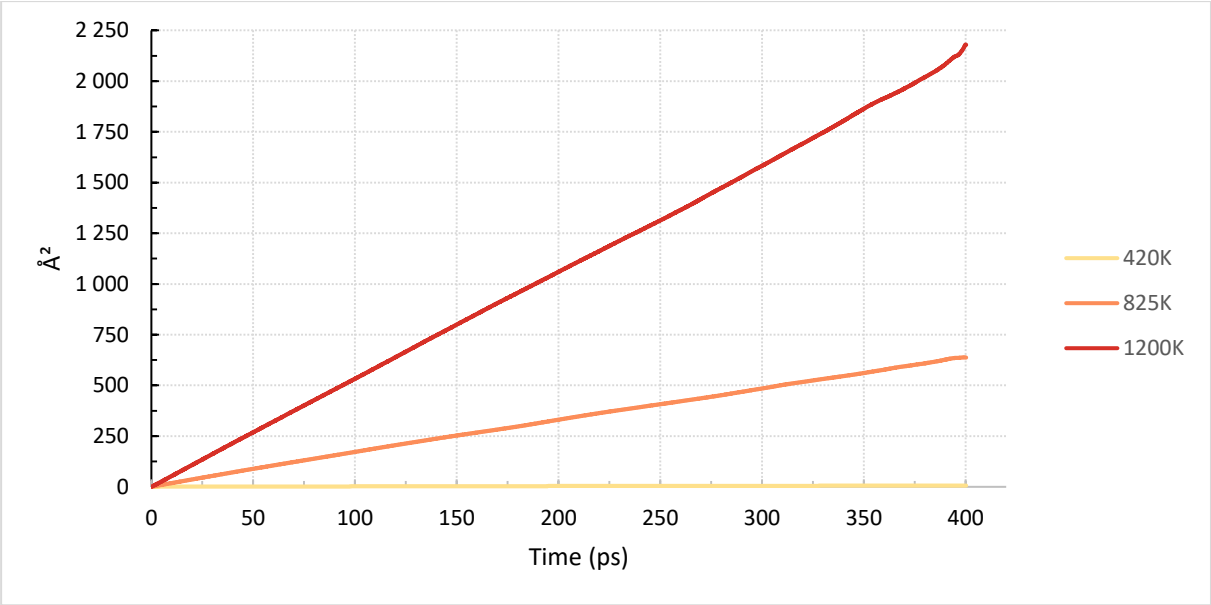

Figure 36 Mean squared displacement of germanium atoms as a function of simulation time in the three different temperatures.

From the calculated mean squared displacements of Sb and Te atoms in the three different temperatures a similar behavior was found for their atomic mobility. Meaning that the movement of both Sb and Te atoms, tends to be slow at the low temperature of 425 K. At the temperature of 825 K, the movement of both these atoms starts to be more moderate, and at the temperature of 1200 K their movement tends to be fast, reflecting their high mobility at the highest temperature studied here.

5.3.2 Diffusion Coefficients

The figure 37 presents the MSD and its linear fit of Ge, Sb and Te atoms at the temperature of 1200 K. From the figure, one can see that at the temperature of 1200 K, Ge is moving the fastest, and Sb and Te are moving similarly. The diffusion coefficients were calculated based on equation 13, presented in chapter 3.4.2, from the slope of the linear regime of the mean squared displacements as a function of simulation time, solving the equations seen in each figure for every element, dividing the x value by six in mathematical calculation software. From the linear fit of the Ge curve the R-squared is 0.9998, for antimony is 0.9996, and for tellurium is 0.9999. At the temperature of 1200 K the linear fits R-squared values are very close to the maximum value of 1. The linear fit of the MSD shows the relationship of the regression model to the simulation data, reflecting to the quality of the fit. R-squared value close to 1 demonstrates good quality of the linear fit.

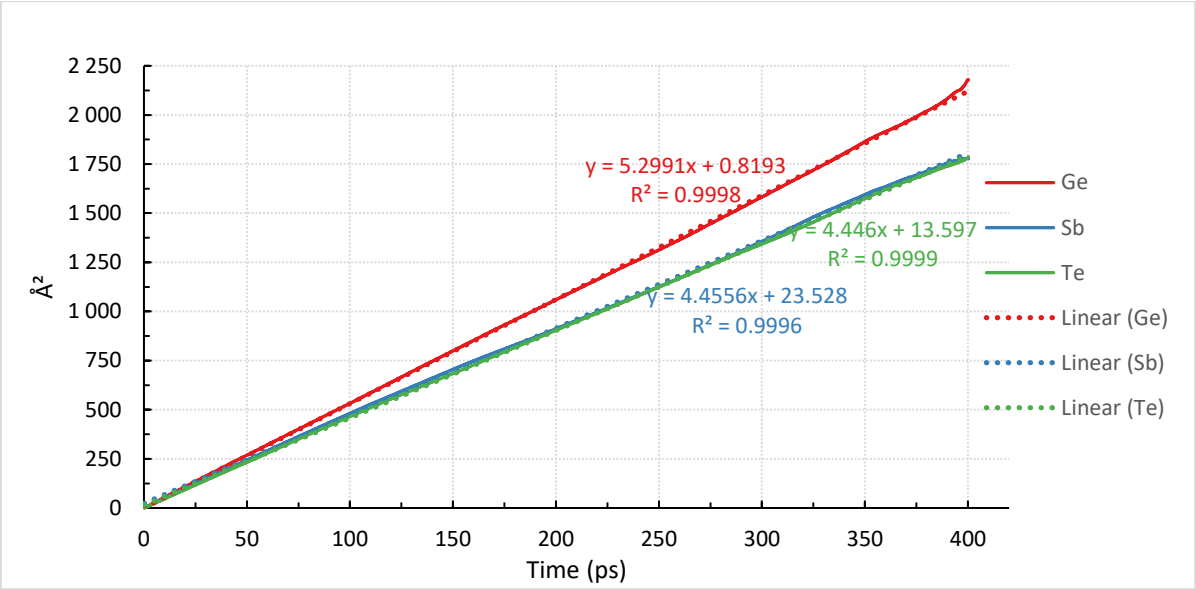

Figure 37 Mean squared displacement and its linear fit for germanium, antimony and tellurium atoms at the temperature of 1200 K.

The figure 38 presents the MSD and its linear fit of Ge, Sb and Te atoms at the temperature of 825 K. From the figure, one can see that at the temperature of 825 K, all the elements are moving in a similar fashion up to the point of 200 picoseconds, while at 400 picoseconds Sb and Te atoms are moving similarly, and Ge atoms are the slowest. The linear fit curve R-squared values are 0.9998 for Sb, 0.9996 for Ge, and 0.9997 for Te. At the temperature of 825 K the linear fits R-squared values are very close to the maximum value of 1.

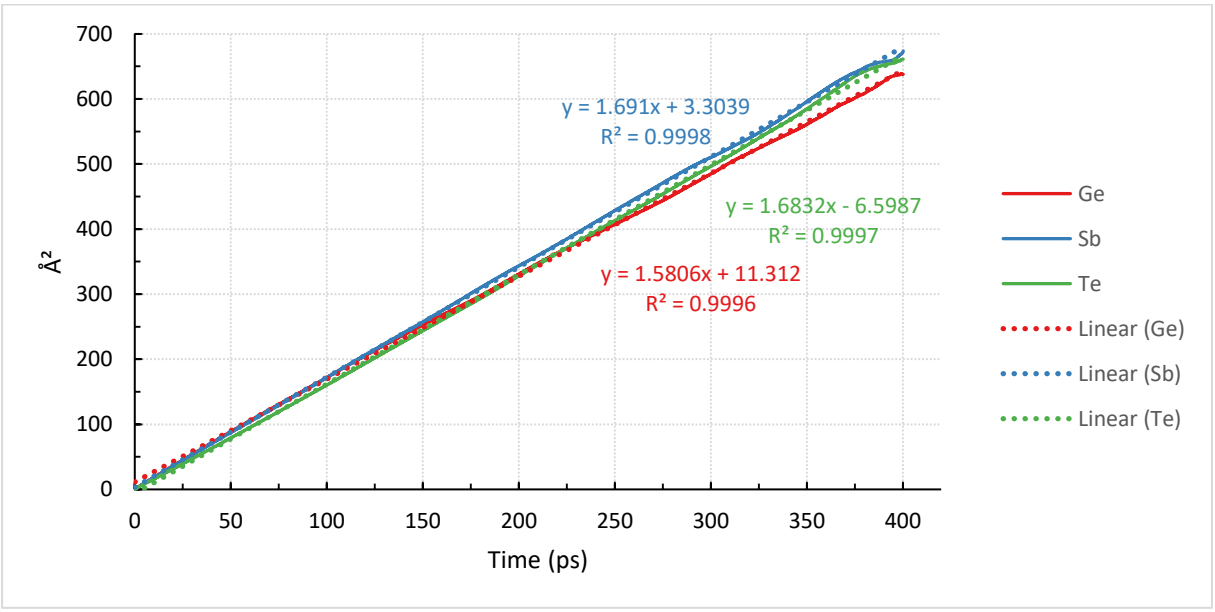

Figure 38 Mean squared displacement and its linear fit for germanium, antimony and tellurium atoms at the temperature of 825 K.

The figure 39 presents the MSD and its linear fit of Ge, Sb and Te atoms at the temperature of 425 K. From the figure, one can see that at the temperature of 425 K, Sb atoms are moving the fastest, Ge atoms are the second fastest, and Te atoms are the slowest. The linear fit curve R-squared values are 0.9992 for Sb, 0.9951 for Ge, and 0.9961 for Te. At the temperature of 425 K the linear fits R-squared values are quite close to the maximum value of 1.

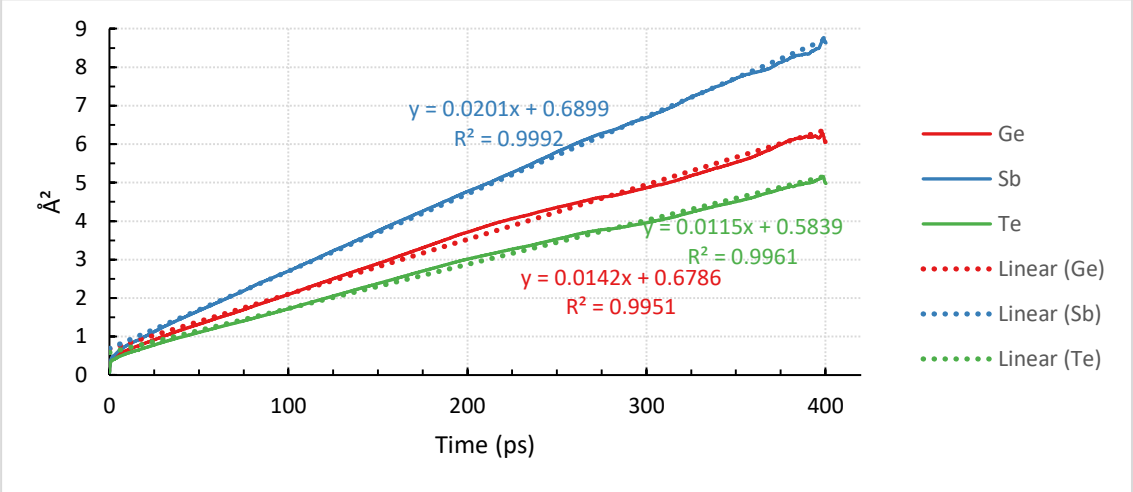

Figure 39 Mean squared displacement and its linear fit for germanium, antimony and tellurium atoms at the temperature of 420 K.

Table 5 presents the diffusion coefficients of Ge, Sb and Te atoms in the three different temperatures studied in this thesis. From the table, one can see that at the temperature of 425 K, the diffusion coefficients for each element are quite similar, being 0.002 for Ge, 0.003 for Sb, and 0.002 for Te. At the temperature of 825 K variations between the diffusion coefficients of the different elements start to occur, obtaining values of 0.263 for Ge, 0.282 for Sb, and 0.281 for Te. At the temperature of 1200 K there are clear differences in the diffusion coefficients of the atomic species, calculated 0.883 for Ge, 0.742 for Sb, and 0.741 for Te. Overall, the calculated values show that in the high (liquid) temperature the three atomics species are moving faster within the simulated structure compared to the low (amorphous) temperature. Also, the diffusion coefficients reflect the dynamical heterogeneity in the atomic mobility of the model, as well as the influence of the supercooled liquid state on the atomic movements of the $Ge_2Sb_2Te_5$ glass structure.

Table 5 Diffusion coefficients of germanium, antimony and tellurium atoms in the three different temperatures.

	Diffusion coefficient		
Atom	420 K	825 K	1200 K
Ge	0.002	0.263	0.883
Sb	0.003	0.282	0.743
Te	0.002	0.281	0.741

6 Discussion and conclusions

The results of this thesis align with previous research performed on Ge-Sb-Te phase-change memory materials. In other simulation studies, similar observations have been reported regarding the radial distribution functions, coordination numbers, bond angles, and local atomic environments inside the $Ge_2Sb_2Te_5$ structure. In addition, the calculated values of mean squared displacements and diffusion coefficients of the atomic species are in good agreement with previously reported values.

Comparing the interatomic distances of different atom pairs at the high (liquid) temperature, calculated values of 2.71 Å for Ge-Ge, 2.82 Å for Ge-Te, 3.0 Å for Sb-Sb and 3.05 Å for Sb-Te have been reported by El Kheir et al. [34] with molecular-dynamics simulations at a temperature of 990 K, using a different ML interatomic potential for amorphous $Ge_2Sb_2Te_5$. These results are in excellent agreement with the respective calculated values in this study at a temperature of 1200 K, being 2.74 Å for Ge-Ge, 2.82 Å for Ge-Te, 3.0 Å for Sb-Sb and 3.0 Å for Sb-Te. Moreover, the coordination numbers distributions calculated in this study at the liquid temperature are in very good agreement with the values reported in Ref. [34] for Ge, Sb, and Te atoms. Overall, from the calculated results presented in this thesis, it was found that the lower the temperature the higher the coordination number of the elements tends to be inside the simulated structure.

Regarding the bond-angle distributions of the Te-Ge-Te and Te-Sb-Te triatomic environments, peaks at around 90 degrees and 169 degrees have been previously reported with quantum-mechanical simulations for amorphous $Ge_2Sb_2Te_5$ at 300 K [8]. These are in very good agreement with the calculated values in this study for the simulations performed at the temperature of 420 K. Overall, it was found here that the bond angle distributions in the model structure become narrower as the simulated temperature decreases, indicating a correlation between the different temperature regimes studied in this thesis (amorphous, supercooled liquid, and liquid) and the atomic environments that exist inside the $Ge_2Sb_2Te_5$ model.

Combining the information from the bond-angle distributions and coordination numbers analysis, the diverse local atomic environments present in the $Ge_2Sb_2Te_5$ material were described consistently in the simulations performed in this thesis. Tetrahedral, trigonal bipyramidal, see-saw and defective-octahedral shaped geometries are representative local

atomic environments suggested from the calculated results in this thesis. These observations are in agreement with previous simulation studies on Ge-Sb-Te materials [7,8,34].

In a previous simulation study with molecular-dynamics simulations at 600 K, based on quantum mechanics, it was found that the mean squared displacements for Ge, Sb and Te atoms were around 15 $Å^2$ after 100 ps from the beginning of the simulation [35]. In this thesis, the mean squared displacement for each atomic species in the $Ge_2Sb_2Te_5$ structure was varying between 2 - 3 $Å^2$ at the temperature of 420 K, and between 150 - 170 $Å^2$ at the temperature of 825 K, considering for both simulations the first 100 ps of the molecular-dynamics run. Taking into account that the temperatures studied in this thesis are in between the temperature used in Ref. [35], the calculated results reported here are in the right range of atomic motion. Wang et al. [36], reported that, at 600 K, the diffusion coefficients were 0.011 for Ge atoms, 0.009 for Sb atoms and 0.008 for Te atoms in the simulated $Ge_2Sb_2Te_5$ structure. These values are in good agreement with the calculated diffusion coefficients of the different elements in the corresponding temperatures studied this thesis.

Overall, regarding the atomic motion in the simulated $Ge_2Sb_2Te_5$ structure the higher the temperature the faster the atomic mobility, for each element, inside the structure. The calculated diffusion coefficients reflect the dynamical heterogeneity in the atomic mobility of the model, as well as the influence of the supercooled liquid state on the atomic movements of the $Ge_2Sb_2Te_5$ glass structure.

7 Outlook

Potential future directions associated with the research performed in this study could include different types of analysis for the atomic structure and dynamical properties, as well as more molecular-dynamics simulations at different target temperatures.

(i) Analysis of the medium-range order structure of the $Ge_2Sb_2Te_5$ model at the three temperatures studied here. Such study could be realized by calculating the distributions of rings and voids, in the simulated structure, and make comparisons between the different temperature regimes (amorphous, supercooled liquid, and liquid).

(ii) Analysis of the intermediate scattering function and calculations of the dynamical propensity for the three different temperatures simulated in this thesis. Such analysis could offer useful insights regarding the dynamical heterogeneity in the simulated $Ge_2Sb_2Te_5$ structure.

(iii) Performing molecular-dynamics simulations for more temperatures within the three temperature regions will provide the opportunity to sample more information across the whole thermodynamical region between the liquid and amorphous states of the simulated $Ge_2Sb_2Te_5$ structure. In that way a better understanding of the atomic mobility and local atomic structure of the material could be obtained.

References

[1] M. Wuttig, N. Yamada, Phase-change materials for rewriteable data storage, Nat Mater 6 (2007) 824–832. https://doi.org/10.1038/nmat2009.

[2] S. Raoux, W. Wełnic, D. Ielmini, Phase Change Materials and Their Application to Nonvolatile Memories, Chem Rev 110 (2010) 240–267. https://doi.org/10.1021/cr900040x.

[3] S. Raoux, Phase Change Materials, Annu Rev Mater Res 39 (2009) 25–48. https://doi.org/10.1146/annurev-matsci-082908-145405.

[4] G.S. Syed, M. Le Gallo, A. Sebastian, Phase-Change Memory for In-Memory Computing, Chem Rev 125 (2025) 5163–5194. https://doi.org/10.1021/acs.chemrev.4c00670.

[5] S.R. Nandakumar, M. Le Gallo, I. Boybat, B. Rajendran, A. Sebastian, E. Eleftheriou, A phase-change memory model for neuromorphic computing, J Appl Phys 124 (2018) 152135. https://doi.org/10.1063/1.5042408.

[6] G.C. Sosso, J. Behler, M. Bernasconi, Atomic mobility in the overheated amorphous GeTe compound for phase change memories, Physica Status Solidi (a) 213 (2016) 329–334. https://doi.org/10.1002/pssa.201532378.

[7] K. Konstantinou, J. Mavračić, F.C. Mocanu, S.R. Elliott, Simulation of Phase-Change-Memory and Thermoelectric Materials using Machine-Learned Interatomic Potentials: Sb2Te3, Physica Status Solidi (b) 258 (2021) 2000416. https://doi.org/10.1002/pssb.202000416.

[8] K. Konstantinou, S.R. Elliott, J. Akola, Inherent electron and hole trapping in amorphous phase-change memory materials: Ge2Sb2Te5, J Mater Chem C 10 (2022) 6744–6753. https://doi.org/10.1039/D2TC00486K.

[9] JR. William D. Callister, David G. Rethwisch, Materials Science and Engineering: An introduction, 10th ed., Wiley, 2018.

[10] A. Redaelli, Phase Change Memory, 1st ed., Springer Cham, Cham, 2018. https://doi.org/10.1007/978-3-319-69053-7.

[11] N. Amini, J. Pries, Y. Cheng, C. Persch, M. Wuttig, M. Stolpe, S. Wei, Thermodynamics and kinetics of glassy and liquid phase-change materials, Mater Sci Semicond Process 135 (2021) 106094. https://doi.org/10.1016/j.mssp.2021.106094.

[12] K. Konstantinou, Computational modelling of structural, dynamical and electronic properties of multicomponent silicate glasses. PhD thesis, University College London (UCL), 2017.

[13] P. Singh, R. M.Singari, R.S. Mishra, A review of study on modeling and simulation of additive manufacturing processes, Mater Today Proc 56 (2022) 3594–3603. https://doi.org/10.1016/j.matpr.2021.12.057.

[14] D.C. Rapaport, The Art of Molecular Dynamics Simulation, Cambridge University Press, Cambridge, 2004. https://doi.org/10.1017/CBO9780511816581.

[15] B.J. Leimkuhler, S. Reich, R.D. Skeel, Integration Methods for Molecular Dynamics, in: 1996: pp. 161–185. https://doi.org/10.1007/978-1-4612-4066-2_10.

[16] S. Sharma, P. Kumar, R. Chandra, Introduction to Molecular Dynamics, in: Molecular Dynamics Simulation of Nanocomposites Using BIOVIA Materials Studio, Lammps and Gromacs, Elsevier, 2019: pp. 1–38. https://doi.org/10.1016/B978-0-12-816954-4.00001-2.

[17] R.S. Katiyar, P.K. Jha, Molecular simulations in drug delivery: Opportunities and challenges, WIREs Computational Molecular Science 8 (2018) 18. https://doi.org/10.1002/wcms.1358.

[18] LAMMPS Molecular Dynamics Simulator, (n.d.). https://www.lammps.org (accessed March 3, 2025).

[19] J. Kemppainen, J.R. Gissinger, S. Gowtham, G.M. Odegard, LUNAR: Automated Input Generation and Analysis for Reactive LAMMPS Simulations, J Chem Inf Model 64 (2024) 5108–5126. https://doi.org/10.1021/acs.jcim.4c00730.

[20] V.L. Deringer, M.A. Caro, G. Csányi, Machine Learning Interatomic Potentials as Emerging Tools for Materials Science, Vol 31, Issue 46, Advanced Materials 31 (2019) 1902765. https://doi.org/10.1002/adma.201902765.

[21] S. Urata, M. Bertani, A. Pedone, Applications of machine-learning interatomic potentials for modeling ceramics, glass, and electrolytes: A review, Journal of the American Ceramic Society 107 (2024) 7665–7691. https://doi.org/10.1111/jace.19934.

[22] S. Klawohn, J.P. Darby, J.R. Kermode, G. Csányi, M.A. Caro, A.P. Bartók, Gaussian approximation potentials: Theory, software implementation and application examples, J Chem Phys 159 (2023) 174108. https://doi.org/10.1063/5.0160898.

[23] J. Behler, Perspective: Machine learning potentials for atomistic simulations, J Chem Phys 145 (2016) 170901. https://doi.org/10.1063/1.4966192.

[24] A.P. Bartók, Gaussian Approximation Potential: an interatomic potential derived from first principles Quantum Mechanics, PhD Thesis, University of Cambridge, 2010.

[25] Central Michigan University, Dr. Sebastien Le Roux, Pr. Valeri Petkov, I.S.A.A.C.S. Interactive Structure Analysis of Amorphous and Crystalline Systems. The radial distribution functions: definitions, (n.d.). https://isaacs.sourceforge.io/phys/rdfs.html (accessed March 25, 2025).

[26] S.J. Grabowski, Triel bond and coordination of triel centres – Comparison with hydrogen bond interaction, Coord Chem Rev 407 (2020) 213171. https://doi.org/10.1016/j.ccr.2019.213171.

[27] Central Michigan University, Dr. Sebastien Le Roux, Pr. Valeri Petkov, I.S.A.A.C.S. Interactive Structure Analysis of Amorphous and Crystalline Systems. Bond properties, (n.d.). https://isaacs.sourceforge.io/phys/bond.html (accessed March 25, 2025).

[28] J. Hafner, BOND-ANGLE DISTRIBUTION FUNCTIONS IN METALLIC GLASSES, Le Journal de Physique Colloques 46 (1985) C9-69-C9-78. https://doi.org/10.1051/jphyscol:1985908.

[29] L.-Y. Xu, Y. Alrefaei, Y.-S. Wang, J.-G. Dai, Recent advances in molecular dynamics simulation of the N-A-S-H geopolymer system: modeling, structural analysis, and dynamics, Constr Build Mater 276 (2021) 122196. https://doi.org/10.1016/j.conbuildmat.2020.122196.

[30] Y. Chen, J. Wang, D.R. Flanagan, Fundamental of Diffusion and Dissolution, in: Developing Solid Oral Dosage Forms, Elsevier, 2017: pp. 253–270. https://doi.org/10.1016/B978-0-12-802447-8.00009-1.

[31] F.C. Mocanu, K. Konstantinou, T.H. Lee, N. Bernstein, V.L. Deringer, G. Csányi, S.R. Elliott, Modeling the Phase-Change Memory Material, Ge2Sb2Te5, with a Machine-Learned Interatomic Potential, J Phys Chem B 122 (2018) 8998–9006. https://doi.org/10.1021/acs.jpcb.8b06476.

[32] CSC – IT Center for Science, (n.d.). https://csc.fi/en/ (accessed May 20, 2025).

[33] W. Humphrey, A. Dalke, K. Schulten, VMD - Visual Molecular Dynamics, J. Molec. Graphics, (1996) 33–38. http://www.ks.uiuc.edu/Research/vmd/.

[34] O. Abou El Kheir, L. Bonati, M. Parrinello, M. Bernasconi, Unraveling the crystallization kinetics of the Ge2Sb2Te5 phase change compound with a machine-learned interatomic potential, NPJ Comput Mater 10 (2024) 33. https://doi.org/10.1038/s41524-024-01217-6.

[35] J. Kalikka, J. Akola, R.O. Jones, Simulation of crystallization in Ge2Sb2Te5: A memory effect in the canonical phase-change material, Phys Rev B 90 (2014) 184109. https://doi.org/10.1103/PhysRevB.90.184109.

[36] R. Wang, Z. Song, W. Song, T. Xin, S. Lv, S. Song, J. Liu, Phase-change memory based on matched Ge-Te, Sb-Te, and In-Te: octahedrons: Improved electrical performances and robust thermal stability, InfoMat 3 (2021) 1008–1015. https://doi.org/10.1002/inf2.12233.

Licensing and permissions