University of Milano - Bicocca
Department of Materials Science
PhD School in Nanostructures and Nanotechnologies
XXV cycle

A Neural Network Potential for the Phase Change Material GeTe: Large Scale Molecular Dynamics Simulations with close to *ab initio* accuracy

A.A. 2011/2012

Ph.D. dissertation of
Gabriele Cesare Sosso

Supervisor
Prof. Marco Bernasconi

To my father, who took me into the mountains

Contents

Introduction

Phase change materials based on chalcogenide alloys are attracting an increasing interest worldwide due to their ability to undergo reversible and fast transitions between the amorphous and crystalline phases upon heating [1]. This property is exploited in rewritable optical media (CD, DVD, Blu-Ray Discs) and electronic nonvolatile memories of new concept, the Phase Change Memories (PCM) [2, 3]. The strong optical and electronic contrast between the crystal and the amorphous allows discriminating between the two phases that correspond to the two bits of binary information zero and one. The material of choice for applications is the ternary compound $Ge_2Sb_2Te_5$ (GST). However, the related binary alloy GeTe has also been thoroughly investigated because of its higher crystallization temperature and better data retention at high temperature with respect to GST.

PCM devices, born thanks to the work of Ovshinsky in the late 1960s [4], offer extremely fast programming, extended cycling endurance, good reliability and inexpensive, easy integration. A PCM is essentially a resistor of a thin film of the chalcogenide alloy with a low field resistance that changes by several orders of magnitude across the phase change. In memory operations, cell read out is performed at low bias. Programming the memory requires instead a relatively large current to heat up the chalcogenide and induce the phase change, either the melting of the crystal and subsequent amorphization (RESET) or the recrystallization of the amorphous (SET).

In the last few years, atomistic simulations based on density functional theory (DFT) have provided useful insights into the properties of phase change materials [5–8]. However, several key issues such as the thermal conductivity at the nanoscale or the origin of the fast crystallization, just to name a few, are presently beyond the reach of ab initio simulations due to the high computational cost. In fact, first principles simulations can deal at most with 10^2 atoms on the timescale of 10^2 ps, while many properties like thermal conductivity or direct simulations of the crystallization process require at least 10^3 atoms on the timescale of 10^3 ps.

The development of reliable classical interatomic potentials is a possible route to overcome the limitations in system size and time scale of ab initio molecular dynamics. However, traditional approaches based on the fitting of simple functional forms for the interatomic potentials are very challenging due to the complexity of the chemical bonding in the crystalline and amorphous phases revealed by the ab initio simulations. A possible solution has been proposed recently by Behler and Parrinello [9], who developed high-dimensional interatomic potentials with close to ab initio accuracy employing artificial neural networks (NN).

In this thesis work, we developed a classical interatomic potential for the bulk phases of GeTe employing this NN technique. The potential was validated by comparing results on the structural and dynamical properties of liquid, amorphous, and crystalline GeTe derived from NN-based simulations with the ab initio data obtained here and in previous works [10]. The NN potential displays an accuracy close to that of the underlying DFT framework at a much reduced computational load that scales linearly with the size of the system. It allows us to simulate several thousands of atoms for tens of ns, which is well beyond present-day capabilities of DFT MD. The development of a reliable interatomic potential with close to DFT accuracy is thus a breakthrough in the modeling of phase change materials.

To date, we employed our NNP in order to investigate three issues:

1. the thermal conductivity of the amorphous phase
2. the viscosity and atomic mobility in the supercooled liquid and overheated amorphous phases
3. the dynamics of homogeneous crystallization of the liquid and amorphous phases

Thermal conductivity is a key property for the PCM operation, as the phase changes corresponding to the memory writing/erasing processes strongly depend upon heat dissipation and transport. Moreover, thermal cross-talks among the different bits is a crucial reliability issue in PCM. Although experimental data on thermal conductivity are available for few materials in this class [11, 12] it is unclear whether the value measured in the bulk phase could also describe the behavior of the material in nanoscaled devices (10-20 nm) which might be smaller than the phonon mean free path. This is particularly relevant for PCM architectures based on nanostructures employing nanowires, colloidal nanoparticles, thin bridges and nanotubes [13, 14]. In fact, amorphous materials can also display propagating phonons with mean free path as long as 0.5 μm. This has been demonstrated for amorphous Si [15] where propagating modes with long mean free path contribute to half of the total thermal conductivity. It is therefore of great technological relevance to assess whether similar propagating modes with long mean free path might be present in amorphous phase change materials as well. Atomistic simulations can provide crucial insights into the thermal transport properties of phase change materials suitable to aid a reliable modeling of the device operation. However, the calculation of the thermal conductivity in an amorphous system requires very long simulations (on the ns scale) of large models (thousands of atoms) that are presently beyond the reach of fully DFT simulations. The use of NN potentials allowed us to compute the mean free path of phonons in a-GeTe and to assess that ballistic transport is inhibited by disorder even at the nanoscale.

The key property that makes some chalcogenide alloys suitable for applications in PCM is the high speed of the transformation which leads to full crystallization on the time scale of 10-100 ns upon Joule heating. What makes some compounds alloys so special in this respect and so different from most amorphous semiconductors is, however, still a matter of debate. The driving force for crystallization of the supercooled liquid is actually the free energy difference between the crystal and the supercooled liquid. However, the crystallization is controlled both by the driving force and by the atomic mobility. The driving force vanishes at melting and increases upon cooling. A large atomic mobility at high supercooling can thus boost the crystallization speed. These conditions can actually be met by fragile liquids. In fact, supercooled liquids are classified as fragile or strong on the basis of the temperature dependence of their viscosity [16]. An ideal strong liquid shows an Arrhenius behavior of the viscosity η from the melting temperature T_m down to the glass transition temperature T_g. On the contrary, in a fragile liquid η is very low down to a crossover temperature T$*$, below which a steep rise of the viscosity (and thus a steep decrease in the mobility) is observed. If T^* is sufficiently far from the melting temperature (T_m), high supercooling and large atomic mobility can be met at the same time. The question is thus whether phase change materials are fragile liquids or not. Due to the high crystallization speed it is unfortunately not possible to measure η below T_m experimentally. We have thus addressed this problem by MD simulations and we have demonstrated that indeed GeTe is a very fragile liquid (fragility index $\sim$ 100). Moreover a breakdown of the Stokes-Einstein relation between the viscosity and the diffusion coefficient is found, which further increase the atomic mobility down to temperatures very close to T_g. Hysteretic effects in the heating of the amorphous phase above T_g have been addressed as well.

Finally, we have performed direct simulations of the homogeneous crystallization of the supercooled liquid and amorphous phases with 4000-atom models and simulation times of several ns. Although similar simulations have been previously performed by fully DFT simulations [17], the limitations in size (about 200 atoms) and time scale (500 ps) prevented a reliable estimate of the size of the critical nucleus and of the crystallization speed, which are instead accessible by our large scale simulations.

This thesis is organized as follows: in the introductory chapter 1 I provide essential information about phase change materials and phase change memories. Chapter 2 is dedicated to theory and methods, while chapters 3,4,5 and 6 are devoted to the results on the thermal conductivity of the amorphous phase of GeTe, on the properties of its supercooled liquid phase and on the homogeneous crystallization from the melt and the overheated amorphous phase.

Chapter 1

Phase Change Materials

1.1 Phase change change memories

Phase change materials are systems of great technological interest as they are nowadays widely employed in optical memories (e.g. DVD-RW) and in a novel and promising class of non volatile memories (NVM) known as Phase Change Memory (PCM) [3, 11]. Both applications rely on the fast and reversible transition between the crystalline and the amorphous phase, induced by heating produced by laser pulses in optical memories and Joule effect in PCM. Crystalline and amorphous phases show large differences in reflectivity and resistivity, exploited in optical memories and PCM for information reading and writing.

Non volatile memories based on phase change memories have been proposed by Ovshinsky [4] around 1960, but they gained importance as effective industrial applications only recently, mainly because of scaling problems suffered by the current NVM Flash technology (*Floating-Gate*, FG [18]). The first PCM devices landed on the market few years ago, and in June 2012 Micron [1] announced the production of devices at the 45 nm scale.

Phase change materials used in optical and electronic devices are typically tellurium based chalcogenide alloys. The family of the pseudobinary compounds $(\mathrm{GeTe})_x(\mathrm{Sb_2Te_3})_y$ of which phase diagram is shown in Fig. 1.1, represents a prototypical system.

So far, the material of choice for PCM is $\mathrm{Ge_2Sb_2Te_5}$ (GST), which shows high transformation speed and great stability of the amorphous phase. Several other compositions along the pseudobinary line have been proposed and scrutinized. Moreover, systems containing Ag, In or Se have been considered as candidates for high temperatures applications [20]. The binary compound GeTe itself shares many structural, vibrational and electronic properties with ternary GST, and in the last few years it has been considered as an alternative to GST because of its higher crystallization temperature.

---

[1] Micron (former Numonyx) is a microeletronics company involved in a fruitful collaboration with the research group I worked with during my PhD.

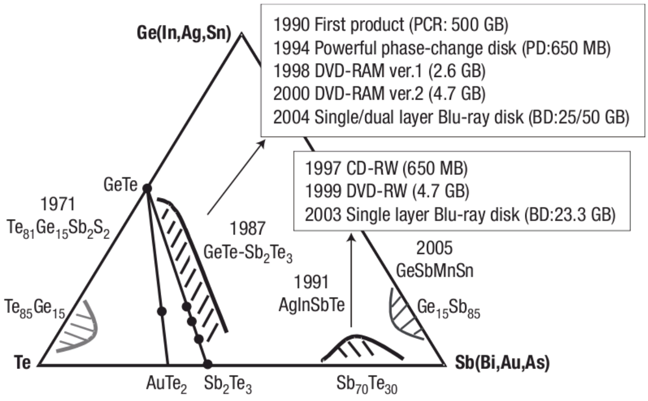

Figure 1.1: Sketch of the Ge Sb Te ternary phase diagram. Composition intervals of interest for applications are highlighted (Ref. [19]).

At the moment different architectures for a PCM device have been proposed (see Fig, 1.2).

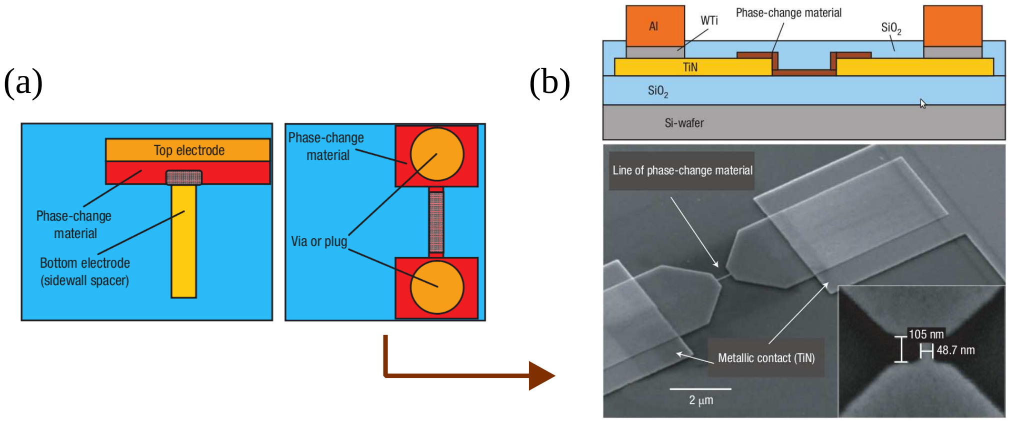

Figure 1.2: Architectures of a PCM cell: a) mushroom cell on the left and and line cell on the right. A SEM image of a line cell from Ref. [19] is shown in panel b).

The architecture of the device commercialized by Micron consists of a resistor made of a thin film of chalcogenide in between a metallic contact and a resistive electrode (mushroom cell, see Fig. 1.2). Memory programming is realized by applying current pulses to the cell, and involves two different processes. In the RESET process the material switches from the crystalline to the (insulating) amorphous phase, while in the SET process the chalcogenide switches back from the amorphous to the crystalline phase. During the RESET process (see Fig. 1.3) a short ($\sim 100\ ns$) current pulse is applied to the crystal, whose temperature in the neighborhood of the resistive electrode rises beyond the melting point (998 K for GeTe [21]) due to Joule effect. After the current pulse the system experiences a fast ($\sim 50\ ns$) cooling by which the liquid freezes into the amorphous phase, characterized by a resistance of the order of $M\Omega$. The process of SET (Fig. 1.3) is achieved by current pulses of length similar to those used in the RESET process but less intense, sufficient to heat the chalcogenide at a temperature below the melting point but high enough to induce the re-crystallization of the material in about 100 ns. The resistance of the cell in this state is of the order or $k\Omega$.

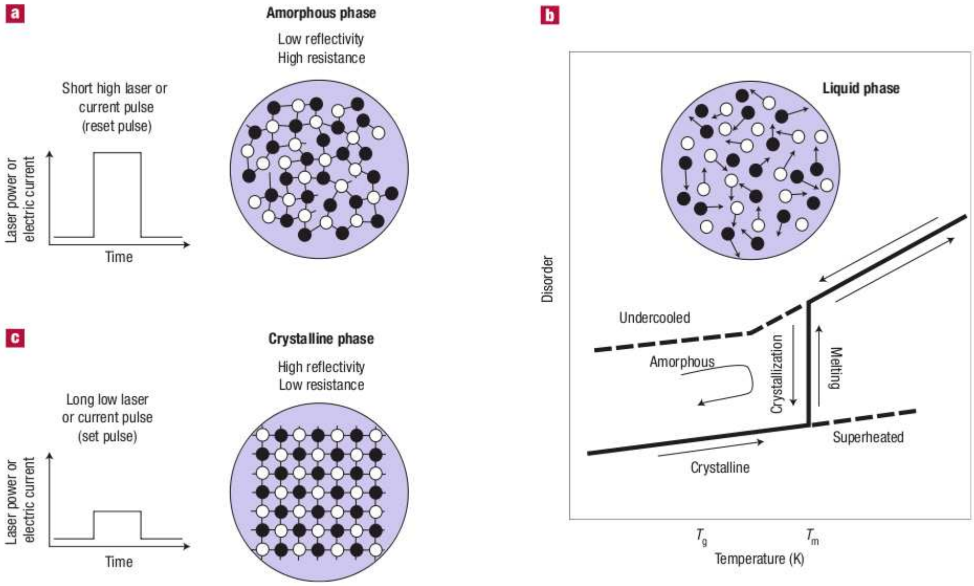

Figure 1.3: RESET (a,b) and SET (c) process typically used in PCM programming. The same concepts apply to the programming of optical memories based on phase change materials like e.g. DVD-RAM and Blu-Ray disks.

It turns out that phase change materials are particularly suitable to be employed in PCM devices, because of the huge resistivity contrast between the amorphous and the crystalline phase (three orders of magnitude) and the amazing speed by which these systems crystallize. Crystallization requires heating by Joule effect of the amorphous phase, which is highly resistive. This can be achieved at low biases in the device thanks to the fact that the amorphous phase displays two different states, each characterized by a different electronic resistance. Let us consider again the SET process (see Fig. 1.4). At the beginning, by applying low biases in the amorphous phase, the system has a high resistance that prevents heat dissipation in the material. However, above a threshold voltage of about one V, the material undergoes a purely electronic transition (threshold switching) that brings the system in a state of much lower resistivity. Current rises consequently in the amorphous region of the cell, dissipating enough heat to reach the crystallization temperature of the chalcogenide. PCM programming is performed on very short timescales (50-100 ns), far shorter than those of the FG (Floating Gate) technology currently employed in NVM. The reading process is also very fast (~ 50 ns), and consists of the measure of chalcogenide resistance at low bias (100 mV). PCM display a promising scalability, as the device operation depends only on the contrast between the properties of the two phases, amorphous and crystalline. In fact, there are indications that this contrast could survive in an active volume of about 10 nm.

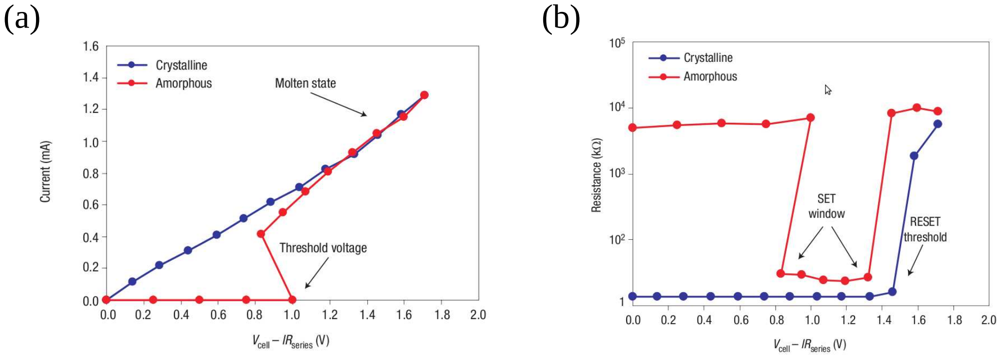

Figure 1.4: Electrical characteristics of a PCM cell (line cell architecture). a) Current as a function of applied voltage during 50 ns pulses and b) evolution of the resistance as a function of voltage after the above mentioned pulses during SET and RESET processes (Ref. [19]).

1.2 GeTe

In Fig. 1.5 the phase diagram of GeTe is shown. In this thesis work, we deal with the stoichiometric composition GeTe. It should be noticed that this system has an eutectic composition at $Ge_{15}Te_{85}$, and that there are more than just one crystalline phase stable at ambient pressure depending on composition and temperature.

At the stoichiometric composition and at room temperature the stable crystalline phase is the so called α phase, which is a trigonal crystal that can be seen as a rocksalt structure distorted along the $< 111 >$ direction (see Fig. 1.7). The octahedral atomic environment of the rocksalt is distorted as well, so that there is an alternance of three short and three long bonds, each atom having three short and three longer bonds (see Fig. 1.6).

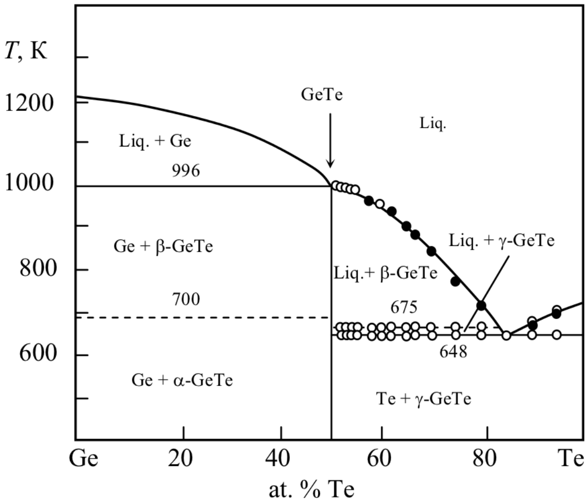

Figure 1.5: Phase diagram of the system Ge-Te [22].

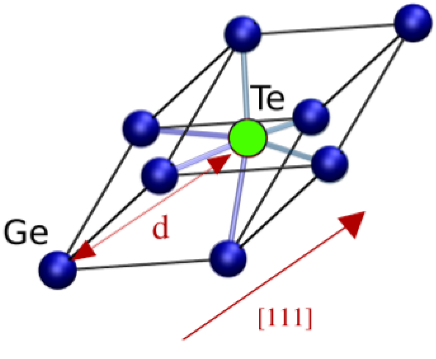

Figure 1.6: Trigonal structure of crystalline GeTe, which can be seen as a cubic rocksalt geometry with an elongation of the [111] diagonal. Ge atoms are on the vertex of the cell while Te atom stands in the center. The distance between Ge and Te atoms along the [111] direction is indicated by *d*.

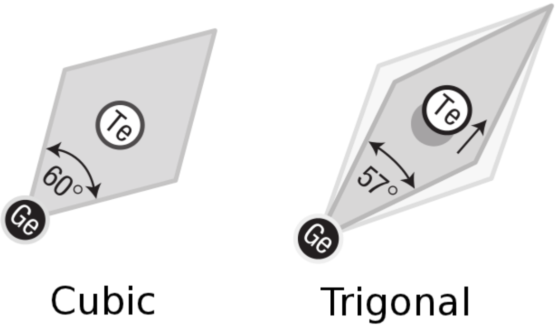

Figure 1.7: Trigonal distortion (α phase) of the cubic phase (β phase) of GeTe. The rombohedral primitive cell of the rocksalt structure of β GeTe is elongated along the $< 111 >$ cubic direction.

α-GeTe is stable up to 700 K, where it undergoes a ferroelectric/paraelectric transition into a rocksalt structure (β-GeTe).

If we keep on heating, we reach the melting point at about 1000 K. In contrast to most tetrahedrally coordinated semiconductors, melting in GeTe is not associated with a semiconductor → metal transition. The conductivity of GeTe upon melting is only 9% higher than that of the solid, classifying GeTe as a semiconductor [23]. The coordination number of liquid GeTe was measured to be 5.1 [24], noticeably larger then the value of 4 common for semiconducting II-VI liquids but clearly smaller than the coordination number 6 found for the metallic IV and III-IV liquids. Raty and coworkers [25] suggested that the alternation of short and long bonds also seen as a Peierls distortion of the cubic phase survives in the local order of liquid GeTe.

The theoretical structure factor of liquid GeTe obtained via DF simulations by Akola *e*t al. [26] is shown in Fig. 1.8 together with experimental data [27]. The agreement is remarkable, and provides a direct proof of the ability of PBE-GGA DFT simulations in describing the structure of phase change materials.

If we cool down the liquid (quenching) below the melting point sufficiently rapidly to avoid crystallization, at a certain temperature known as glass transition temperature T_g we obtain a glass.

The structural properties of amorphous GeTe, and in general of the amorphous phase of the other GeSbTe systems, have been matter of debate for years, together with the mechanism driving the phase transitions. A step toward comprehension of these issues was provided by extended x-ray absorption fine structure (EXAFS) experiments [29–31], which showed that the local structure of GST and GeTe changes upon amorphization. The average coordination number of Ge atoms in a-GST and a-GeTe decreases from

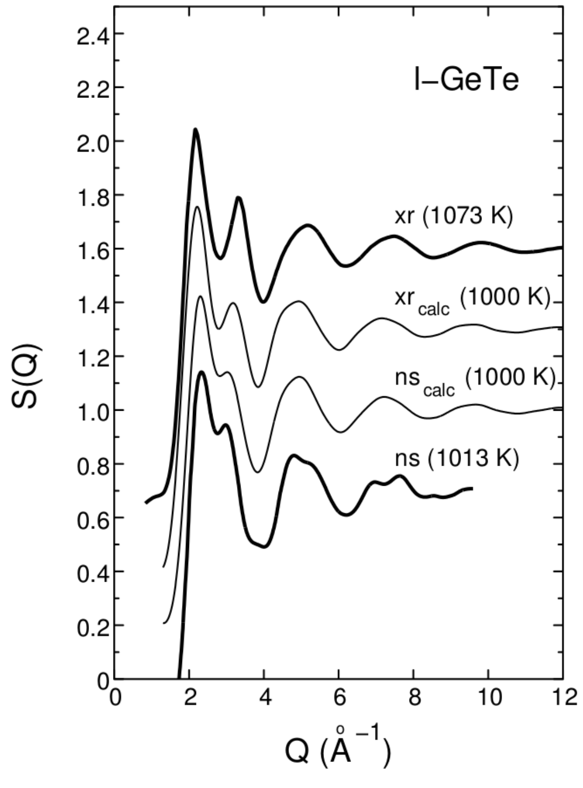

Figure 1.8: Structure factor of liquid GeTe calculated by Fourier transforming the calculated partial pair correlation functions with weights appropriate to x-ray (xr) and neutron scattering (ns). Experimental curves are shown: xr and ns at 653 K from Ref. 27, ns at 773 K from Ref. 28.

six-fold in the crystal to fourfold in the amorphous phase. Models of amorphous GeTe and GST have been obtained by quenching from the melt in DFT molecular dynamics simulations [3, 32, 33], revealing that indeed Ge atoms are mostly fourfold coordinated, but as opposed to the proposal inferred from EXAFS data [29] only approximately one third of Ge atoms display tetrahedral coordination, whereas the majority of Ge and all Te and Sb atoms are in a defective octrahedral environment (see Fig. 1.9). The fraction of tetrahedra actually depends on the preparation conditions of the amorphous. In fact, it has been shown that models of as-deposited a-GST generated by DFT simulations [34] show a larger fraction of tetrahedra than models generated by quenching from the melt.

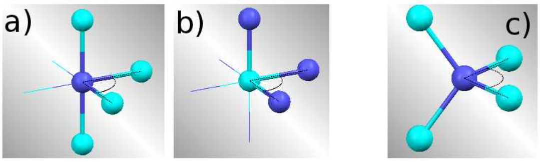

Figure 1.9: Different atomic environment in a-GeTe. a), b) Defective octahedral environments, c) tetrahedral environment.

In Fig. 1.10 the experimental structure factor $S(Q)$ of a-GST is compared with the $S(Q)$ obtained by DFT simulations [8]. The agreement is remarkable, and it can be improved by reverse Monte Carlo refinement [35]. Signatures of Ge in tetrahedral sites were assigned to specific features of the Raman spectrum of a-GeTe [36]. Based on ab initio calculations, Mazzarello *et. al* [10] provided an assignment of the Raman spectrum of GeTe to vibrations of specific local structures in the amorphous network which, by comparison with experimental Raman spectra (see Fig. 1.11), yields a compelling evidence of the existence of both tetrahedral and defective octahedral structures in a-GeTe.

First principles simulations provided useful insights also into the crystallization mechanism of the amorphous phase. In fact, it has been shown [26] that the amorphous network of a-GeTe and a-GST contains a consistent number of four membered rings (see Fig. 1.12 and Fig. 1.9) that are the building blocks of the crystal phase itself.

The realignment of the squared rings to form the crystalline phase was suggested to be aid by the presence of nanovoids in amorphous phase (see Fig. 1.13 and section 2.3.4 as well). The role of four membered rings has been underlined also by studying the amorphization of GST under pressure [37]. By pressurizing the cubic crystalline phase, first Te atoms move to fill the vacancy voids giving rise to topological defects in the rocksalt crystal made of homopolar Ge-Sb bonds in squared rings rotated by 45 degrees with respect to the crystalline axis (see Fig. 1.14). By further increasing pressure, the concentration of the latter topological defects increases up to 21 GPa at which the system transforms into an amorphous phase. The amorphization occurs on the time scale of a few picoseconds, and it is driven by the higher compressibility of the amorphous phase with

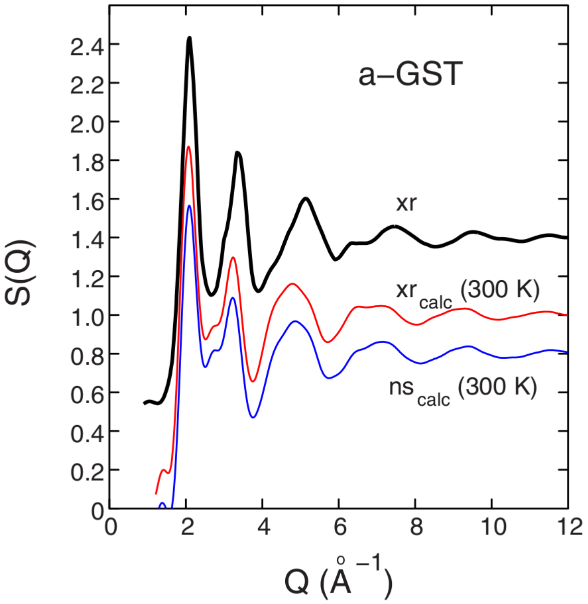

Figure 1.10: Structure factor of a-GST. The experimental curve is displaced by 0.4. From Ref. 8 and Ref. 27.

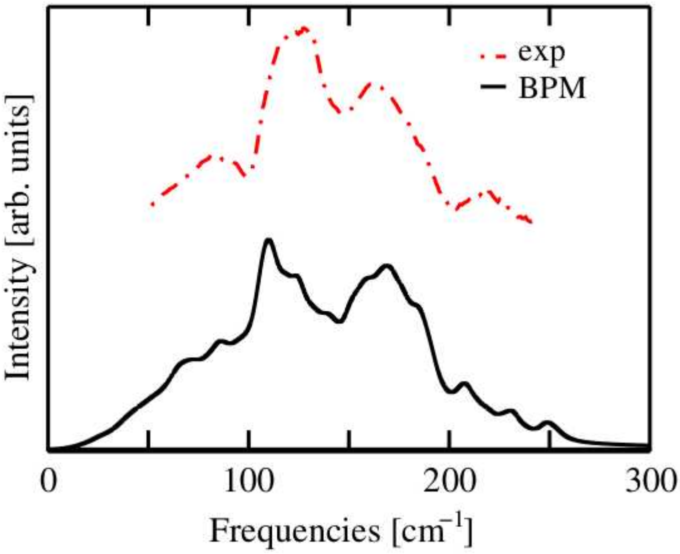

Figure 1.11: Reduced Raman spectrum of a 216-atoms a-GeTe model from Ref. 10 compared with the analogous experimental spectrum from Ref. 36.

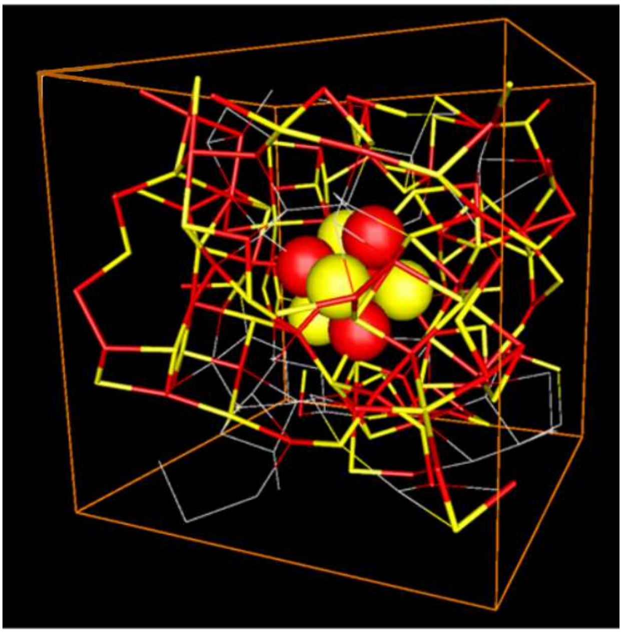

Figure 1.12: Simulation box of a-GeTe (216 atoms) with four membered rings highlighted. Red, Ge; yellow, Te. From Ref. 26.

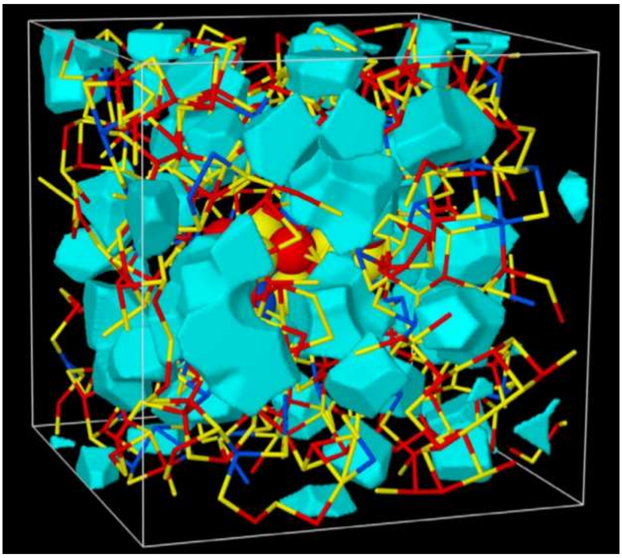

Figure 1.13: Cavities in melt-quenched GST (8,2,11). The volumes have been constructed around regions at least 2.8 Å away from any atom.

respect to the crystal [37]. Direct molecular dynamics simulations of crystallization of GST and GeTe have been recently [17] performed, since the timescale on which the phase transitions take place for this class of materials is short, of the order of few hundred of ns. However, due to the still high computational cost the results reported in Ref. 17 refer to small models (10^2 atoms), where the effect of periodic boundary conditions strongly affects the simulations. Overall the simulations performed on GST, GeTe and related materials demonstrated the reliability of the DFT-PBE framework in describing the structural and dynamical properties of materials in this class.

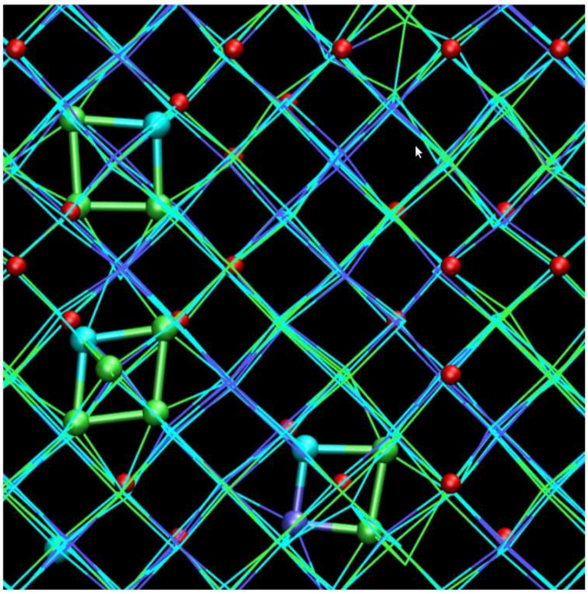

Figure 1.14: Snapshot of a molecular dynamics simulation at 14 GPa showing the appearance of squared rings formed by homopolar Ge-Sb bonds rotated by 45 degrees with respect to the crystal axis. Large spheres indicate atoms involved in the formation of rotated rings (blue: Ge; green: Sb; light blue: Te). Small red spheres indicate the position of Ge=Sb vacancies. From Ref. 37.

1.3 Thermal transport

Thermal transport plays an important role in PCM. In fact, the very idea upon which the devices are based is the phase transition between the crystalline and the amorphous phase and viceversa driven by Joule heating. It is therefore clear that heat dissipation and transport can strongly affect the performances of the device. The electrothermal modeling of the device requires a value for the thermal conductivity of the chalcogenide which has been obtained from measuring in thick films. However, the transferability to the nanoscale device of the values of the thermal conductivity measured in the bulk needs further validation. Due to the small size of the device, ballistic transport might in fact occur even in the amorphous phase. This issue will be even more critical in the novel PCM architectures based on nanowires [13] or colloidal nanoparticles [14]. Actually, it has been shown e.g. for amorphous Si [15] that even disordered systems can display propagating vibrations with a very long mean free path, up to 0.5 μm, thus possibly greater than the size of the nanodevice itself. Unraveling the microscopic origin of thermal conductivity in amorphous phase change materials is therefore needed to assess whether ballistic transport might be observed in nanodevices or not. To this aim, one has to consider the different kinds of vibrations that can be observed in disordered systems. Owing to the structural disorder, one can anticipate the existence of non-harmonic motions, usually described by two-level states (TLS) and tunneling [38]. However, in this thesis we limit ourselves to harmonic and quasi harmonic vibrations, neglecting TLS and tunneling which are expected to play a role only at low temperatures. Plane wave acoustic phonons are present for wavelengths sufficiently long to consider the amorphous as a continuum. Up to which wavevector value such a description is still meaningful is a central question. On the other hand, the structural units can collectively execute their own vibrations corresponding to optical branches. In glasses, something quite analogous to optic modes should be expected at sufficiently small $\vec{q}$. Propagating plane waves are surely not a description of vibrations above the Ioffe-Regel crossover frequency ω_{co} [39], at which their mean free path is comparable to the wavelength. A new taxonomy has thus been introduced [40] to address properly the different kinds of vibrational modes in glasses. The vibrational eigenmodes are called vibrons rather than phonons, as the latter refer to plane waves like propagating modes. The extended vibrons are called extendons. Each one of these is distributed rather evenly over the entire system. In other words, the inverse participation ratio (IPR) of the jth extendon, defined as

$$IPR = \frac{\sum_i \left| \frac{\vec{e}(j,i)}{\sqrt{M_i}} \right|^4}{\left(\sum_i \frac{|\vec{e}(j,i)|^2}{M_i} \right)^2}. \tag{1.1}$$

is basically of the order of $1/N$. In Eq. 1.1 $\vec{e}(j,i)$ are phonon eigenvectors and the sum over i runs over the N atoms with masses M_i in the unit cell. According to this definition, the value of IPR varies from $1/N$ for a completely delocalized phonon, to one for a mode completely localized on a single atom. The extendons can be either propagative ($\omega < \omega_{co}$) or diffusive ($\omega > \omega_{co}$). The formers are called propagons and the latter diffusons. As ω is raised further, diffusons reach a mobility edge beyond which the modes become truly localized (the IPR is low). These modes are then called locons. The eigenvectors of a given locon decay exponentially away from its center, as electrons do in Anderson localization [41]. In real systems, these locons are usually at very high frequencies, and they usually display a rather small contribution to the total density of vibrational states. The taxonomy of vibrations in glasses (from Ref .42) is shown in the scheme of Fig. 1.15

and is used in Fig. 1.16 to classify the phonons of amorphous Si in Ref .42.

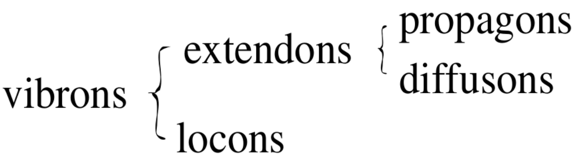

Figure 1.15: Taxonomy of vibrations in glasses.

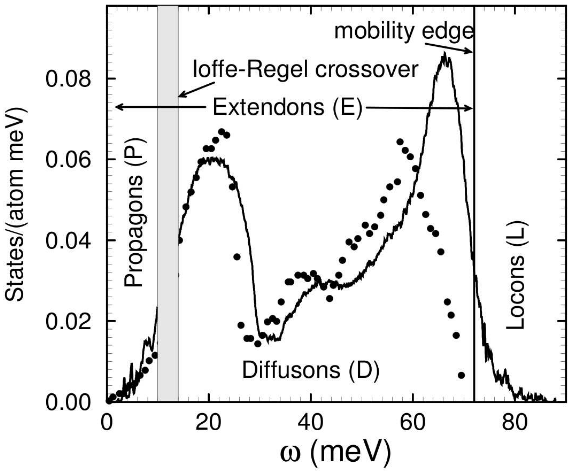

Figure 1.16: Phonon density of states for a 4096-atom model of amorphous Silicon compared with experimental neutron scattering data from the work of Allen *et al.* [42]. The different spectral regions correspond to different type of phonons.

In these thesis work, we investigated the microscopic origin of thermal conductivity as the sum of two different contributions: the first one comes from propagating vibrations, that can be computed by means of the Boltzmann Transport Equation, while the second accounts for diffusons treated according to the theory developed by Allen and Feldman [40]. Both approaches are discussed in section 4. These approaches follow the work by Donadio *et al*, who in the last few years investigated the origin of thermal conductivity in amorphous silicon [15] showing that indeed both propagating vibrations and diffusons contribute to the thermal conductivity.

1.4 Supercooled liquid phase and crystallization

Let us consider a liquid at a certain temperature above its melting temperature T_m. As the melt is quenched below the melting point, the system becomes increasingly rigid. Eventually, when the temperature is brought low enough, a liquid-solid transition takes place (with some exceptions like liquid helium). A liquid may solidify in two ways: discontinuously to a crystalline solid or continuously to an amorphous solid (glass). In a quenching experiment carried out at a sufficiently low cooling rate, crystallization is usually the route taken to reach the solid state. But at sufficiently high cooling rates, many materials follow a different behavior. The crystallization is bypassed, and the liquid phase persists until a lower temperature, the glass transition temperature T_g, is reached. Here the second solidification scenario is realized. The atomic mobilities D become too small for structural rearrangements required to reach thermal equilibrium. Thus, the undercooled liquid is "frozen in" an amorphous solid. It is worth noticing that T_g depends upon the cooling rate, T_g shifting to lower temperatures when the cooling rate decreases The glass transition temperature is commonly defined as the temperature at which the viscosity equals $1 \cdot 10^{12}$ Pa·s. The viscosity η of the liquid is linked to the self diffusion coefficient D via

$$D = \frac{k_B T}{6\pi\eta r_{sph}} \tag{1.2}$$

where r_{sph} is the radius of the particles, usually taken as the average of the atomic Van der Waals radii if for a multicomponent system. Whether a glass can be formed at low cooling rates (*easy or good glass former*) or whether this may only be accomplished via rapid quenching (*marginal or bad glass former*) as in the case of phase change materials, it depends on the atomic mobility and on the driving force for crystallization. Equivalently, it also The temperature dependence of η varies among materials, classifying them in strong or fragile supercooled liquids.

As sketched in Fig. 1.17, for a strong liquid, the viscosity follows an Arrhenius behavior from T_m down to T_g while in the case of a fragile liquid the viscosity follows an Arrhenius behavior only above a crossover temperature $T^\dagger$ where the curvature of $\eta(T)$ displays a marked change and below which a super Arrhenius behavior is observed and usually described by the so called Vogel-Tammann-Fulcher (VTF) function[16]

$$\eta = \eta_0 e^{\frac{E}{k_B(T-T_0)}} \tag{1.3}$$

with η_0, E and T_0 as fitting parameters, even though many other functional forms have been proposed. Given that the viscosity at both T_g and T_m would be fixed, crystallization upon cooling of a strong undercooled melt is more likely bypassed even at moderate cooling rates as compared to the case where the melt is a fragile liquid. Hence, a link between fragility and the ease of glass formation can be drawn. Glass transition temperatures can experimentally be assigned via differential scanning calorimetry (DSC) since the transition

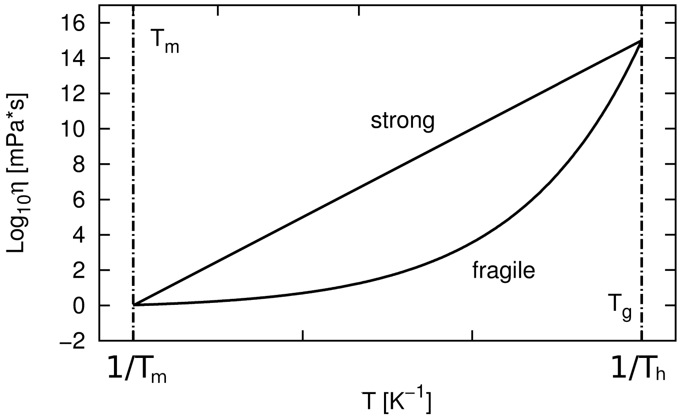

Figure 1.17: Schematic Angell plot for strong and a fragile supercooled liquids showing the temperature dependence of viscosity between the melting (T_m) and the glass transition temperature (T_g).

from the glass to an undercooled liquid appears as an endothermic peak upon cooling. However, in the case of phase change materials such a measurement is difficult, because the system crystallizes so fast that the glass transition is easily obscured by crystallization. T_g is thus often considered equal to the crystallization temperature T_c obtained upon heating with a moderate heating rate. To date, no reliable values if T_g for phase change materials are available from experiments.

The crystallization process takes place via nucleation and growth. The nucleation can be homogeneous when it takes place in the bulk of the material or heterogeneous when it occurs at the interface with the surrounding material. Nucleation and growth are usually investigated by means of classical nucleation theory [43, 44] (CNT). In this theory crystal nucleation is controlled by the competition between the free energy gain due to the liquid-crystal transformation and the free energy loss associated with the formation of the liquid-crystal interface. Assuming that the crystal nucleus is a sphere or radius R, its total free energy cost is

$$\Delta G = -\frac{4}{3}\pi R^3 n_s \Delta\mu + 4\pi R^2 \gamma \tag{1.4}$$

where n_s is the number density of particles in the solid, $\Delta\mu$ is the difference between the liquid and the solid chemical potentials, and γ is the liquid-solid interfacial tension. This ΔG gets trough a maximum at $R_c = 2\gamma/(n_s\Delta\mu)$ (critical nucleus size) and the height of the free energy barrier is given by

$$\Delta G^c = \frac{16\pi\gamma^3}{3(n_s\Delta\mu)^2} \tag{1.5}$$

CNT predicts a us the steady state nucleation rate I_{ss}, that is the frequency by which crystal nuclei pop up at a certain temperature as

$$I_{ss} = s_c\Gamma_z\frac{6D}{\lambda^2}e^{-\left[\frac{\Delta G^c}{k_BT}\right]} \tag{1.6}$$

where s_c is the number of surface sites in the critical cluster, λ is a typical jump distance in the diffusion process and Γ_z is the Zeldovich factor, typically of the order of 1/100 and smoothly dependent on temperature. The factor governing I_{ss} are thus self diffusion coefficient, and $\Delta\mu$ which enter in the definition of ΔG_c. Upon cooling the driving force for crystallization ($\Delta\mu$) displays an opposite trend with respect to atomic mobility. This is sketched in Fig. 1.18: the driving force increases upon cooling while the diffusion coefficient rises up.

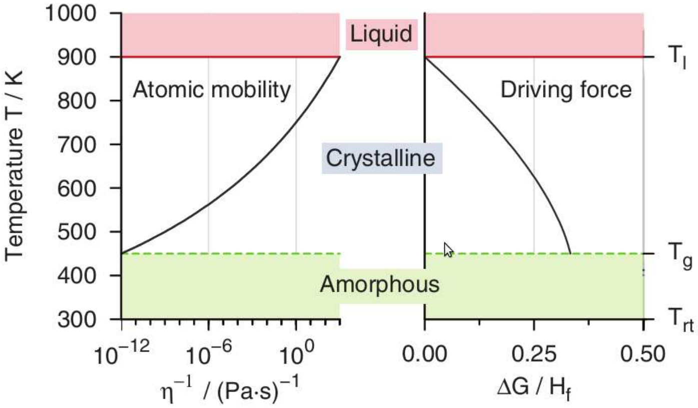

Figure 1.18: Schematic representation of the opposite trend upon cooling of atomic mobility with respect to the driving force for crystallization. From Ref. 5.

One of the most noticeable issues of CNT is the fact that the nuclei are supposed to be spherical. This is not true in most cases, and indeed there are corrections to CNT that can take into account the effect of strain and the fact that the nucleus can have nonspherical shapes with a diffuse interface with the liquid [45, 46].

Since it has been established [47] that crystal growth in phase change materials is a diffusion-limited process, following Ediger *et al* [48] we can assume continuous growth of supercritical nuclei with crystal growth velocity U given by

$$U(T) = U_{kin}(T)\left[1 - exp\left(-\frac{\Delta\mu(T)}{RT}\right)\right] \tag{1.7}$$

where $\Delta\mu(T)$ is the driving force for crystallization and $U_{kin}(T)$ is a kinetic coefficient that includes geometrical factors and -more importantly- the temperature dependence of the diffusion coefficient, that usually can be described by an Arrhenius behavior. $U_{kin}(T)$ is usually written as

$$U_{kin}(T) = \gamma_s \frac{6D(T)}{\lambda} \tag{1.8}$$

where γ_s is a geometric factor which value is typically in the range 0-1. It is possible [49] to estimate $\Delta\mu(T)$ from the latent heat of melting ΔH_m and the melting temperature T_m. For glass forming system, and particularly for fragile chalcogenide liquids, the following expression can be used [50]:

$$\Delta\mu(T) = \frac{\Delta H_m (T_m - T)}{T_m} \left(\frac{2T}{T_m + T} \right) \tag{1.9}$$

The temperatures at which nucleation and growth exhibit their respective peak values are not the same, and also one process might generally be dominant over the other. In fact, one can distinguish between nucleation-dominated and growth-dominated crystallization (see Fig. 1.19).

In the case of nucleation dominated crystallization, the time needed to crystallize an amorphous or a liquid bit does not depend on its volume. Growth-dominated materials, on the other hand, do exhibit a volume dependence: the smaller the volume, the shorter is the time for complete the crystallization. One can distinguish between three temperature regimes relevant for the crystallization kinetics. The first regime is located around the glass transition temperature where crystallization is ought to proceed very slowly due to the low atomic mobility. The slow crystallization enables the experimental direct observation of nucleation and growth [51, 52]. A second temperature regime of interest is situated just below the liquid temperature where crystallization is slow because of the small driving force for crystallization. thermal analysis. The experimentally most challenging yet technologically relevant temperature regime is located somewhere between liquid and glass transition temperature, where crystallization proceeds the fastest.

Recently, Orava *et al.* [53] succeeded in measuring the crystallization speed $U(T)$ in the temperature range of interest for device application. They unravel the crystallization kinetics by employing differential scanning calorimetry (DSC), which measures the energy required to heat the sample. They utilize an ultrafast-heating calorimeter that enables heating rates higher than $4{\cdot}10^4$ K$\cdot$ s^{-1} to investigate the crystallization of GST. The high heating rates enabled them to study crystallization at temperatures ranging from 450 to 650 K of interest for device operation. On the basis of several assumptions on the crystallization process, the crystal-growth velocity and then U_{kin} from Eq. 1.7 are extracted. The calculated values of D are however inconsistent with the values of η expected at T_g. Therefore a fraction SER has been proposed with

$$U_{kin} \propto D \sim \eta^{-\xi} \tag{1.10}$$

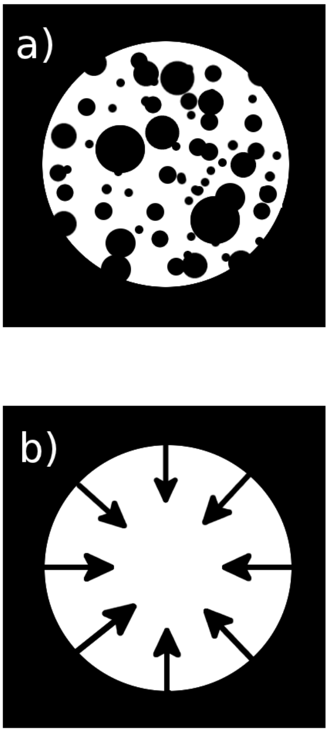

Figure 1.19: Schematic representation of a) nucleation dominated crystallization and b) growth dominated crystallization in which the crystal grows from the crystal/liquid(amorphous) interface.

with ξ less than one. The dependence of U_{kin} and thus η on temperature also suggested that GST is a fragile liquid (see Fig. 1.20). However, a number of crucial assumptions have been made in this work: for example, the authors ignore the contribution of nucleation to their DSC data, and the glass transition temperature is set to 383 K with no direct measurements.

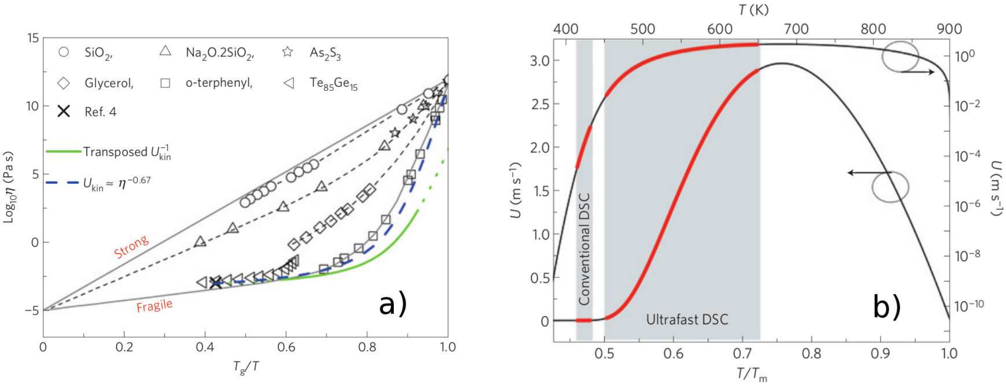

Figure 1.20: a) Angell plot for temperature dependence of viscosity. The solid green line shows the data for U_{kin}^{-1} in GST. The dashed green lines are extrapolations based on fitting to Eq. 5.6. b) Crystal growth rate of GST as a function of temperature. Both of the panels are adapted from Ref. [53].

However, direct experimental measurements of viscosity are not possible in the case of phase change materials because of their fast crystallization. In this thesis work we investigated the dynamical properties of supercooled liquid GeTe by calculating independently the diffusion coefficient and the viscosity. The analysis of this two quantities as a function of temperature provided direct evidence of the fragility of GeTe (and possibly of other phase change materials) and the breakdown of SER close to T_g.

Chapter 2

Methods

2.1 Neural Networks

Neural Networks (NN) constitute a broad subject, for they concern many different fields in which they can play different roles. In order to provide an useful -and possibly clear-definition of NN in the context of this thesis work, we begin by making a distinction between Biological and Artificial Neural Networks. The formers are the building blocks of our brains, and can be defined as sets of interconnected neurons. It was actually in the attempt of both understand and emulate one of the most powerful and effective machinery that nature has ever developed that Artificial Neural Networks showed up at the beginning of the forties [54]. From then on, they undergo an impressive growth in terms of popularity among many different research areas, so that a comprehensive definition of an Artificial Neural Network cannot be properly formulate. Indeed, there is no such thing as "The Artificial Neural Network", but instead this term includes a whole class of algorithms. In this thesis we will obviously refer only to Artificial Neural Networks -from now on simply Neural Networks, NN- and choose for them the following definition [55]:

> A cognitive information processing structure based upon models of brain functions. In a more formal engineering context: a highly parallel dynamical system with the topology of a directed graph that can carry out information processing by means of its state response to continuous or initial input.

Nowadays NN are extensively employed as valuable tools in classification problems, such as speech [56] or text recognition [57], optimization tasks like solving the traveling salesman problem [58], prediction issues as financial market analysis [59] or weather forecast [60] and fitting algorithms [61]. The latter context is the one in which we are interested in, because it has been proven [62] that in principle it is always possible to construct a NN able to fit to arbitrary accuracy any continuous, real-valued function of any dimension. An interatomic potential is basically a representation -to be precise an approximation-of the potential energy surface (PES) of the system, which in turn is a continuous real valued function, often of high dimensionality.

2.1.1 The curse of dimensionality

NN are not the only way by which one can try to approximate a function. Among the many alternatives, splines [63], Taylor expansions [64], Gaussian based methods [65] can be used. One of the most tricky problem when using most of this approaches, consists in the fact that they often fail when the dimensionality of the function is huge, like for example in the case of an high dimensional PES. This issue is known as "the curse of dimensionality" [66]. In fact, in any real fitting procedure we are always forced to work with a limited quantity of input data, that is, a limited database of known data -actually belonging to the function that we want to fit- that constitute the input of our fitting algorithm. Let us take a function $f\{\vec{x}\}$ that depends on a set of variable $\{\vec{x}\}$ of dimension n. When the dimensionality of the function increases, the volume of the space spanned by $\{\vec{x}\}$ increases exponentially, so that our dataset becomes very sparse, in which case it provides a very poor representation of f. In the case of a PES of a system containing for example one thousand atoms, we would like to fit a function that depends on one thousands three dimensional vectors, that is the set of atomic coordinates. This is a formidable task, that requires extreme flexibility. NN are not that susceptible to the curse of dimensionality as the majority of the above mentioned fitting methods. At first, they can exploit the fact that the input variables are generally correlated in some way, so that the data points do not fill out the entire input space but tend to be restricted to a sub space of lower dimensionality. The input of a NN can be build such that the input variables are parametrized in basis function that can be regularized in order to improve the result. Besides, for most mappings of practical interest, the value of the output variables will not change arbitrarily from one region to another, but will typically vary smoothly as a function of the input variables. NN can be nested to create hierarchical algorithms that provide a very high degree of flexibility, being able to handle fitting procedures of high dimensional functional mapping in a massively parallel way.

2.1.2 Feed-forward network mapping

Feed-forward neural networks provide a general framework for representing non-linear functional mappings between a set of input variables and a set of output variables. This is achieved by representing the non linear function of many variables in terms of compositions of non-linear functions of a single variable, called activation functions. In this case each multivariate function can be represented in terms of a network diagram such that there is a one-to-one correspondence between components of the function and the elements of the diagram, that is, any topology of network diagram can be translated into the corresponding mapping function, as long as the network is a purely feed-forward one. In feed-forward networks, there are no feedback loops, that is, it is possible to attach successive numbers to the inputs and to all of the hidden and output units such that each unit only receives connections from inputs or previous units. In this case, outputs can be always expressed as a deterministic function of the inputs, and so the whole network represents a multivariate non-linear functional mapping. A feed-forward NN (from now on, simply NN), is a mathematical object made of different layers. A very simple NN topology is shown in Fig. 2.1, and for sake of simplicity we will refer to it in the following discussion.

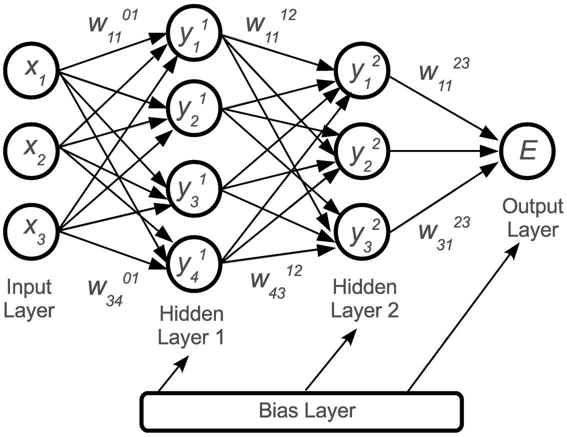

Figure 2.1: Sketch of simple NN topology with two hidden layers.

Each layer is a set of nodes, and each node mimics a single neuron of a biological NN. There are always a single input layer and a single output layer, plus a certain number of so-called hidden layers and optionally a bias layer. The number of hidden layers has to be determined a posteriori. As the term "hidden" suggests, nodes in hidden layers have no physical meaning, being just auxiliary mathematical constructs that provide the required flexibility of the NN. Increasing the number of input layers and/or the number of nodes in each hidden layer increases the number of parameters by which our approximate functional form depends on, so that the flexibility of the NN increases along with the capabilities to fit more complicated functions. Still, increasing flexibility does not always leads to better results. This issue is know as overfitting, that will be addressed in section 2.1.5.

Each node in each layer is connected to the nodes in the adjacent layers by "weights", i.e. the fitting parameters of the NN that define the functional mapping (see Fig. 2.1). Weight connecting node i in layer k with node j in layer l is indicated by w_{ij}^{kl} in Fig. 2.1. In addition to this, nodes in the hidden layers and in the output layers are connected to a bias layer via a set of bias weights b_i^j, which purpose will be discussed in section 2.1.3.

The output of the NN is calculated as follows. At the beginning, each point x_i of the fitting dataset is assigned to a different node in the input layer. The values y_j^1 of the nodes in the first hidden layer are then calculated in two steps: first, for each node a weighted sum χ_j^1 of the input x_i^1 is calculated and the bias weight b_j^1 is added:

$$\chi_j^1 = b_j^1 + \sum_i^4 w_{ij}^{01} \cdot x_i \tag{2.1}$$

this corresponds to a linear combination of the input coordinates using the weights as coefficients. Then, a non-linear function f_j^1 is applied to χ_j^1 in order to allow the NN to fit arbitrary function (activation functions are discussed in detail in section 2.1.3), yielding the value y_j^1 of the hidden node:

$$y_j^1 = f_j^1(\chi_j^1) \tag{2.2}$$

Once all values of all nodes in a hidden layer have been determined, the values of the subsequent layer can be calculated and so forth until the output of the NN is obtained. The complete functional form of the NN depicted in Fig. 2.1 is given by:

$$E = f_1^3 \left(b_1^3 + \sum_{i=1}^{3} w_{i1}^{23} + f_i^2 \left(b_i^2 + \sum_{j=1}^{4} w_{ji}^{12} f_j^1 - \left(b_j^1 + \sum_{k=1}^{3} w_{kj}^{01} x_k \right) \right) \right) \tag{2.3}$$

In principle the NN output can be a vector of values. Here we have chosen to sum up the result of the hidden layers manipulations in a single output node. The ability of the NN to accurately describe a function depends on the numerical values of the weights, that are typically determined in iterative optimization processes (see section 2.1.4).

2.1.3 Activation functions and the bias weights

The activation functions (also called ”transfer functions” or ”basis function”), typically converge asymptotically to a constant value for very large and very small arguments, but in between they display a non-linear region. Possible choices are the sigmoid function

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.4}$$

or the hyperbolic tangent, but also Gaussian, exponential and even periodic functions. Sigmoid, hyperbolic tangent and Gaussian functions converge to a constant number for very small and very large arguments. This property is important for the numerical stability of the NN fitting process. On the other hand, these functions have a narrow range of possible function values, which not necessarily coincides with the range of values of the functional mapping in which we are interested in. Therefore, from the last hidden layer to the output layer, activation functions are usually linear in order to avoid any constraint in the range of output values. By a proper adjustment of the weights, activation functions can be shifted up and down or left and right with respect to the reference frame, and they also can undergo a slope change or a rescaling. In particular the bias weights act as an adjustable offset and shift the nonlinear region of the activation functions. The high flexibility of e.g. the hyperbolic tangent as a function of the weights is shown in Fig. 2.2, and is one of the reasons why NN are capable to adapt very accurately to any kind of function by combining a large number of these simple elements.

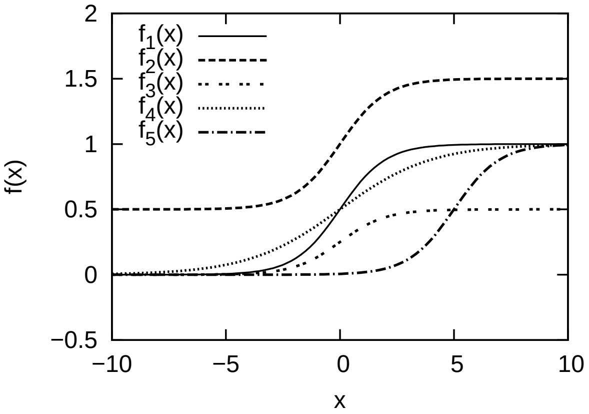

Figure 2.2: Illustration of the flexibility of the sigmoid function. $f_1(x) = \frac{1}{1+e^{-x}}$, $f_2(x) = \left(\frac{1}{1+e^{-x}}\right) + 0.5$, $f_3(x) = \left(\frac{1}{1+e^{-x}}\right) \cdot 0.5$, $f_4(x) = \frac{1}{1+e^{-(x \cdot 0.5)}}$, $f_5(x) = \frac{1}{1+e^{-(x-5)}}$.

2.1.4 Error back-propagation and training

We can define an error function, such as the sum-of-squares errors

$$\Gamma = \frac{1}{2N} \sum_{i=1}^{N} (E_{i,NN} - E_{i,ref})^2 \tag{2.5}$$

where N is the number of point in the dataset, $E_{i,NN}$ is the i-th output of the NN and $E_{i,ref}$ is the reference value of the dataset. If the activation functions of a NN are differentiable, then the output is a differentiable function of both the input variables and the weights (bias weights included) and the error is itself a differentiable function of the weights. We can therefore evaluate the derivatives of the error with respect to the weights, and these derivatives can then be used to find weight values which minimize the error function, by using an optimization method like for example a simple gradient descent. The algorithm for evaluating the derivatives of the error function is known as "back propagation", since it corresponds to a propagation of errors backwards through the NN.

The process by which the weights are iteratively improved until they provide a reasonable approximation of the underlying function is called "training" or "learning", and each iteration of this process is known as "epoch" in the NN context. In spite of the availability of a large number of optimization algorithms, finding low minima in a high-dimensional weight parameter space is a formidable task. Although NN can in principle be used to fit any real-valued function to arbitrary accuracy, in practical fits this knowledge is not of much help, and there is no hope to find the global minimum. According to experience, however, NN do not often get stuck in a poor local minima but tend to find sufficiently accurate local minima in the majority of cases.

2.1.5 Overfitting

Employing an extremely flexible functional form immediately rises the possibility of "overfitting", that is, a set of the points of the dataset is fitted very accurately while other points not included in the training set are poorly described. In other words, overfitting is an improvement of the fit in one region of the configuration space at the cost of a poor quality in another region. A simple example is shown in Fig. 2.3, where the dashed line connects the training points with a very low error. However, a lower error cannot guarantee an improved fitting if the underlying function. In fact, the solid line in Fig. 2.3 represents a fit with an even lower error, since the curve cross exactly each training point, but it is clear that the fit gives very poor result in intermediate regions between the points, including many local extrema not present in the training data.

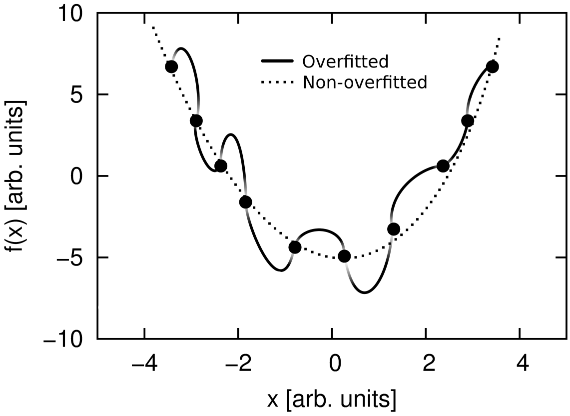

Figure 2.3: A schematic representation of overfitting of a simple function.

This is a typical example of overfitting, an issue particularly relevant in NN as is very easy to increase flexibility of the NN functional form by simply adding more hidden layers (or more nodes in each hidden layer). In a high dimensional weight space, a visual inspection of the fit quality is impossible. A commonly employed method in order to avoid overfitting is the "early-stopping" method, where the input dataset is split into a training set, which is used to update the weights, and a test set, which is not used in the optimization procedure. A comparison of the errors of the training and the test sets allows for an estimate of the properties of the NN fit. Typically, along the different epochs the test error decreases together with the training error until the former reaches a minimum after which it begins to rise up. This means that from then on we are improving the accuracy of the fit with respect the training points on the expense of the regions in configuration space that lie in between them. Thus, in the early stopping method the set of weights that minimizes the test error is usually taken as the best result of the fit.

2.1.6 Neural Network potential energy surfaces for atomistic simulations

NN as a way to construct potential energy surfaces for atomistic simulations have been proposed and implemented in many different fashions. Most of them deal typically with molecular systems, for which a single NN is often used to calculate the total energy the molecule. In this approach the input data are typically the atomic internal coordinates of the molecular system. The use of a single NN for all the atoms is easy to implement, NN standard software is readily available and the number of weights is small, so that fitting turns out to be easy and fast. However, these kinds of NN potentials suffer from severe limitations. At first, they have a fixed dimensionality, i.e., once fitted, the NN potential cannot be applied to systems with a different number of atoms. Further, the number of degrees of freedom of the system must be necessarily small. In practice, single-NN PES are typically used for systems with up to about ten to twenty degrees of freedom corresponding to molecules containing only six or seven atoms. Still, there exists several useful application of single-NN potentials [67–69] Beside isolated molecules in the gas phase, potentials based on a single NN have also been applied to small molecules interacting with surfaces [70–72].

NN potentials became competitive in the field of molecular simulations only if they can provide energies and forces for large systems containing many -i.e. thousands and more- atoms. While most classical force fields and empirical potentials easily fulfill this requirement, for NN potentials it is not immediately obvious how high-dimensional potentials can be constructed. There is in fact a number of conceptual problems that need to be solved to apply NN to the construction of high-dimensional PES. First of all, a NN has a fixed number of input nodes corresponding to the degrees of freedom of the system. Once the weights have been determined, the architecture of the NN cannot be changed. Therefore, a PES based on a standard NN is only valid for a specific system size, as each additional atom would contribute another three degrees of freedom. It is neither desirable nor feasible to construct a separate NN potential for each system size. Further, the number of input nodes of a NN needs to be kept reasonably small. With growing number of input nodes, the NN evaluation becomes computationally more demanding, and the number of weights increases rapidly, slowing down the fitting process.

Thus, in spite of being very accurate and successful, the low-dimensional NN potentials for isolated molecules, or for simple molecular interactions with ideal surfaces discussed above will remain special-purpose solutions. They can very accurately represent the PES of a single molecule or small clusters, but a straightforward extension to larger systems is not possible. In most cases is not even possible to apply these potentials to systems containing several molecules of the same kind, since the intermolecular interactions are completely missing in the NN potentials. The first NN scheme to deal with high-dimensional PES has been suggested already at a very early stage of NN potential development for carbon and the binary C-H system [73]. In this scheme the NN is not directly used to express the energy of the system as a function of the coordinates of the atoms. Still, it is conceptually important as it first develops the idea to employ a NN of variable size, which has later been extended to an efficient high-dimensional NN PES by improving the functional form of the Tersoff potential [74]. The breakthrough in this approach is that the NN is now applicable to systems containing an arbitrary number of atoms, and the problem of variable system size has been solved by decomposing the chemical environment of each bond in the system into a variable number of input vectors characterizing three-atom chains, which all have the same dimensionality. A drawback of the method is that its accuracy is certainly limited, since the functional form of the Tersoff potential is kept and only a few close atoms in the chemical environments of the bonds are taken into account. Starting from this work, Tersoff functional form has been recently abandoned and the energy of the system has been rewritten as the sum of atomic energies [75], derived from individual atomic NN. Therefore, for each atom in the system there is one separate NN, and each NN can process multiple input vectors depending on the local environment of each atom, and the output of each NN is the atomic energy. The method has been successfully applied in the construction of NN potential for silicon [76].

A further improvement with respect to the above mentioned approaches came in 2007 thanks to the work of Behler and Parrinello [9], who developed the method that has been applied in this thesis work to the study of the phase change material GeTe. Again, the total energy of the system is written as the sum of atomic energies:

$$E_{Tot.} = \sum_{i=1}^{N} E_i(\{\vec{r}\}) \tag{2.6}$$

The novelty of the methods lies in the fact that the architecture of the NN is fixed for a given chemical element, so that for each atom a standard NN can be used. Only one input vector of fixed dimensionality is needed per atom to describe its local chemical environment, which is considered up to a certain cutoff radius function that introduce the following cutoff function:

$$f_c(r_{ij}) = \begin{cases} 0.5\left[cos\left(\frac{\pi r_{ij}}{r_c}\right)+1\right] & \text{for } r_{ij} < r_c, \\ 0 & \text{for } r_{ij} > r_c, \end{cases} \tag{2.7}$$

The number of input nodes used to describe an atomic environment is decoupled from the number of neighboring atoms by transforming the Cartesian coordinates of the atoms in the local environment into special types of many-body so-called "symmetry functions ". These symmetry functions provide information on the radial and angular arrangement of neighbors for each atom in the system. A sketch depicting the NN scheme adopted by

Behler and Parrinello is shown in Fig.2.4.

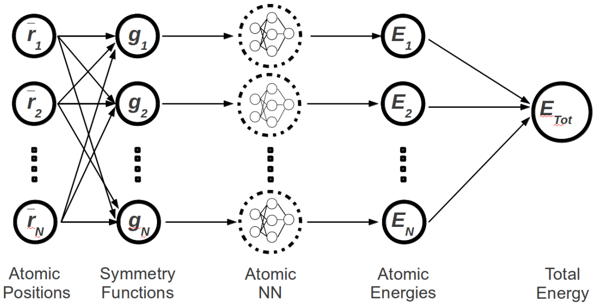

Figure 2.4: Sketch of a NN according to the Behler and Parrinello framework.

A general problem when constructing NN potentials is the correct inclusion of the invariance of the total energy with respect to symmetry operations such as the exchange of two atoms or translations and rotations of the system as a whole. In the scheme proposed by Behler and Parrinello, this feature is realized by a proper choice of the symmetry functions. For a generic system containing N atoms, the coordinates are transformed into a vector of symmetry function values $\{\vec{g}_i\}$. Each symmetry function value depends on all coordinates of the atoms in the environment. The set of values of symmetry functions is the input of an atomic NN which yields an atomic energy. For a given chemical element, the architecture and weights of the atomic NN are constrained to be the same. Therefore, the total energy of the system is invariant with respect to the exchange of the positions of any two atoms of the same element. This NN scheme is applicable to any number of atoms: if an atom is removed or added from the system, we remove or add the corresponding atomic NN. The NN weights can be determined using total energies from electronic structure calculations, because a partitioning into atomic energies before fitting is not required. From now on, we will restrict our discussion to the NN proposed by Behler and Parrinello applied in this thesis to GeTe.

Symmetry Functions

Two different class of symmetry functions have been used: radial and angular symmetry functions. The formers are written as sums of two-body terms, while the latter contain also three-body terms. To describe the radial environment of atom i, two different kind of radial symmetry functions have been employed:

$$G_i^1 = \sum_j f_c(R_{ij})$$
$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}) \tag{2.8}$$

Function G_i^1 is simply the sum of the cutoff functions with respect to all neighboring atoms j, while function G_i^2 is a sum of Gaussians centered at a certain radial distance R_s and multiplied by the cutoff function. These "shifted" G_i^2 functions are suitable to describe a spherical coordination shell around the reference atom. For small values of η and $R_s = 0$ function G_i^2 reduces to function G_i^1. The radial distribution of neighbors can be described by using a set of radial functions with different spatial extensions, e.g. G_i^1 functions with different cutoff radii, or G_i^2 functions with different cutoffs and/or η and R_S parameters. Typical forms of the radial symmetry functions are plotted in Fig. 2.5 for several parameter values.

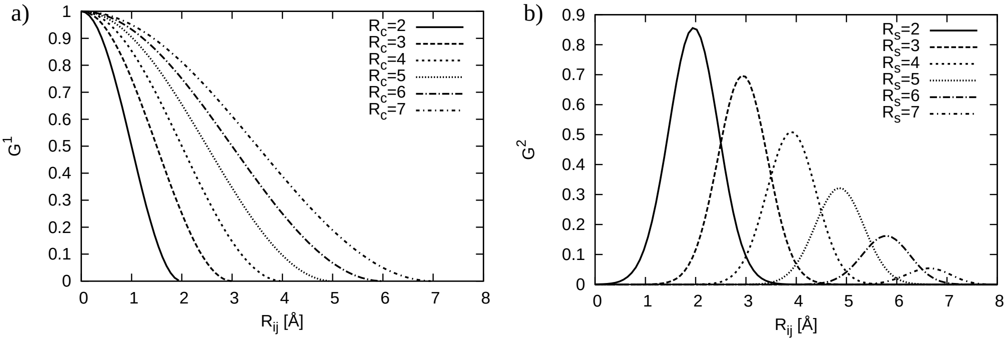

Figure 2.5: Radial symmetry functions. a) G_i^1-type symmetry functions for different cutoff radii. b) G_i^2-type symmetry functions for different radial distances R_s with η=2 and R_c=8 Å respectively.

Since the radial functions are sums over all neighbors, they give a measure of the pair correlation function at various distances. The angular symmetry functions are functions of the angle θ_{ijk} between the central atom i and two neighbors j and k as

$$G_i^3 = 2^{1-\xi} \sum_{j,k\neq i}^{all} (1 + \lambda cos\theta_{ijk})^\xi \cdot e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)^2} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}) \tag{2.9}$$

The parameter λ can have values +1 or -1, shifting the maxima from 0° and 180°. The angular resolution is controlled by the parameter ξ. High ξ values yield a narrower range of nonzero symmetry function values. Therefore, a set of angular functions with different ξ-values can be used to obtain a measure of the bond angle distribution function of each reference atom. The angular distribution is sampled at various distances from the central atom by a suitable choice of η and R_c, which control the radial part. An example of angular symmetry function is depicted in Fig.2.6.

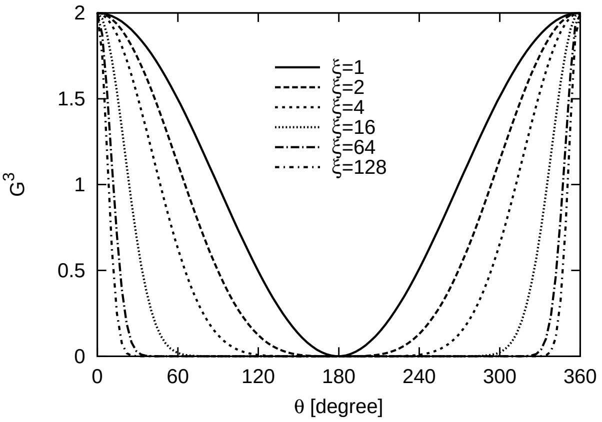

Figure 2.6: Angular symmetry functions G_i^3 for several values of ξ with $\lambda = 1$.

The parameters that define the symmetry functions are fixed in the training process of the NN. The resulting set of function values is different for each atomic environment. Therefore, the total number of symmetry function values describing a given structure is clearly larger than the number of degrees of freedom of the system. This ensures that the full dimensionality of the system is captured, and the resulting redundancy of the information is not usually a problem for a NN algorithm. It should be noted that the introduction of a cutoff radius r_c represents an issue if long-range interactions are present in the system under investigation. However, the Behler and Parrinello NN approach has been successfully extended by the same authors in order to properly include electrostatic interactions, and a NN potential for the strongly ionic compound ZnO has been recently constructed [77]. However, in this thesis work long-range interactions has been neglected.

Forces and stress evaluation

The NN energy is an analytical function of the symmetry functions and thus of the atomic coordinates. Atomic forces and the stress tensor can thus be evaluated analytically. In order to calculate the force component $\vec{F}_k$ acting on atom k is sufficient to apply the chain rule:

$$\begin{aligned}\vec{F}_k &= -\frac{\partial E}{\partial \vec{R}_k} \\ &= -\sum_{i=1}^{N} \frac{\partial E_i}{\partial \vec{R}_k} \\ &= -\sum_{i=1}^{N}\sum_{j=1}^{M_i} \frac{\partial E_i}{\partial G_{i,j}} \frac{\partial G_{i,j}}{\partial \vec{R}_k} \end{aligned} \tag{2.10}$$

where $\vec{R}_k$ is the position of atom k, N is the number of atoms and M_i is the number of symmetry functions describing the environment of atom i. The first term, $\frac{\partial E_i}{\partial G_j}$ is given by the architecture of the NN, the second term $\frac{\partial G_j}{\partial \vec{R}_k}$ is given by the definition of the employed symmetry functions.

Concerning the stress tensor, we refer to its static contribution. Since the energy is a function of interatomic distances $\vec{R}_{kj} = \vec{R}_k - \vec{R}_j$, the stress tensor can be obtained from the virial theorem as [78]

$$\sigma_{\alpha,\beta}^{static} = \sum_{i=1}^{N}\sum_{k=1}^{N} R_{ik,\alpha} \cdot F_{k,\beta} \tag{2.11}$$

where α and β are Cartesian coordinates, i and k are atomic indexes and j runs over the number of symmetry functions. Eq. 2.11 must be evaluated taking into account the fact that $F_{k,\beta}$ come from radial or angular symmetry functions derivatives. It is worth noticing that partitioning the total stress of a system into atomic contribution is in general a non-unique choice. However, in the case of the NN architecture we employed, atomic stress tensors can be easily defined and calculated.

Extrapolation

NN are powerful fitting methods which allow interpolation of the points in the dataset but not extrapolation outside the configurational space spanned. Luckily is pretty easy to rapidly check if we are outside the known configurational space or not. In fact, every local atomic environment is described by a set of symmetry functions, that depending on the atomic environment assume different values. The first step of the NN fitting procedure is the construction of the symmetry function input vector, so that we can keep memory of the minimum and maximum values assumed by each symmetry function for the whole input dataset. If a certain atomic configuration cause one or even more than one symmetry function to assume values outside the range included in th fitting, a so called extrapolation occurred, and we cannot trust our NNP results. Fixing this issue is relatively simple. Once detected the atomic environment that causes the extrapolation, we can add it to the input dataset and fit again, obtaining a NNP that would be now able to behave correctly also for that particular configuration. Still, care must be taken since absence of extrapolations cannot guarantee the reliability of NNP.

2.1.7 The Kalman filter

An optimization scheme which has become very popular in the context of neural networks is the extended Kalman filter. The global extended Kalman filter is a very sophisticated algorithm originating from estimation and control theory [79]. It is used as an online learning method, that is the NN weights are optimized after each individual training point has been taken into account. In order to update the weights $\vec{w}$ after a certain number of dataset points n, an history of previous updated is used, according to the following relations:

$$\begin{aligned}
\bar{\bar{K}}(n) &= \lambda^{-1}\bar{\bar{P}}(n-1)\bar{\bar{J}}(n)[\bar{\bar{I}}+\lambda^{-1}\bar{\bar{J}}^{T}(n)\bar{\bar{P}}(n-1)\bar{\bar{J}}(n)]^{-1} \\
\bar{\bar{P}}(n) &= \lambda^{-1}\bar{\bar{P}}(n-1)-\lambda^{-1}\bar{\bar{K}}(n)\bar{\bar{J}}^{T}(n)\bar{\bar{P}}(n-1) \\
\vec{w}(n) &= \vec{w}(n-1)+\bar{\bar{K}}(n)[\bar{\bar{E}}_{ref}(n)-\bar{\bar{E}}_{NN}(\vec{w}(n-1))] \qquad (2.12)
\end{aligned}$$

where $\bar{\bar{K}}$ is the so called Kalman matrix gain, $\bar{\bar{P}}$ is the covariance matrix[1] , $\bar{\bar{I}}$ is the identity matrix and $\bar{\bar{J}}$ is the Jacobi matrix which elements are defined by:

$$J_i = \frac{\partial E}{\partial w_i} \qquad (2.13)$$

For each training point, a three steps procedure takes place. At first, the Kalman filter gain matrix is updated using the covariance matrix of the previous epoch and the current weight derivatives in the Jacobi matrix. Then the new vector of weight parameters is determined using $\bar{\bar{K}}$. Finally, the covariance matrix is updated according to Eq.2.12. A forgetting schedule is introduced via λ to ensure that only the recent history of updates is taken into account for update of point n:

$$\lambda(n) = \lambda_0\lambda(n-1) + 1 - \lambda_0 \qquad (2.14)$$

The constant λ_0 is usually chosen to be around 0.995. Adapting the weight parameters after each training point is computationally rather costly. In its adaptive form, the Kalman filter is not used to update the weights after each individual training point, but an error threshold α is defined in terms of the actual root mean squared error (RMSE) of the full training set. Only if the error of a training point is larger than the product of α and the current RMSE, the point will be used to update the weights. This can reduce the computational effort significantly. For the construction of NN potentials, the extended Kalman filter often shows a performance which is superior to other optimization

---

[1] In general, $P_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)]$, where $\mu_i = E(x_i)$ is the expected value of the i-th element of the vector $\vec{x}$.

algorithms [80, 81], mainly because is less likely to get trapped in shallow local minima The fitting procedure of a NN employing the Kalman filter can take hours or even days, according to a huge number of variables like the NN architecture, the number and the complexity of the symmetry functions chosen, the dimension of the input dataset and the numerical precision required to the fit. Nevertheless, once the NN potential is constructed it can be used to predict energy and forces in a straightforward way. In principle, the computational load scales linearly with system size.

2.2 Molecular dynamics

Molecular dynamics [78, 82] is a computational technique which provides the time evolution of the positions and momenta of a finite system of atoms, considered as classical particles, once the initial conditions and the interaction potential are specified. Depending on the choice of the interatomic forces there is a trade-off between accuracy and size of the system and time scale affordable by the simulations: on parallel computers available today empirical force fields enable the simulation of millions of atoms for microseconds, whereas more accurate methods which take explicitly into account the electronic structure at the DFT level (e.g. Car-Parrinello method) are limited to hundreds of atoms for few hundreds of picoseconds.

Neglecting quantum effects like zero-point energy and tunneling, which may be relevant only for the lightest atoms, the motion of atoms is accurately described by Newtonian dynamics at low temperatures.

2.2.1 Integration of the equation of motions

Within molecular dynamics the equations of motion are integrated by means of numerical algorithms which rely on a discrete time-scale. The systems considered typically include many atoms interacting in a complex way, thus the trajectories are extremely sensitive to initial conditions: small unavoidable numerical errors lead rapidly to a divergence with respect to the correct trajectories [82]. However, this does not hamper the calculation of observable properties of the system, which usually are time correlation functions, decaying before the divergence takes place, or statistical averages, which rely on the ergodic hypothesis (see section 2.2.2). The velocity Verlet algorithm [83], which we adopt in our simulations, being a democratization of Hamilton's principle of stationary action, has the important property of providing trajectories which are overall very close to real trajectories between the assigned endpoints in space-time, although with different initial velocities than the input ones. According to this algorithm at every timestep Δt the position $\vec{R}_i$ and velocity $\dot{\vec{R}}_i$ of atom i are updated as

$$\begin{aligned}\vec{R}_i(t+\Delta t) &= \vec{R}_i(t)\Delta t + \frac{\vec{F}_i(t)}{2M_i}\Delta t^2 \\ \dot{\vec{R}}_i(t+\Delta t) &= \dot{\vec{R}}_i(t)\Delta t + \frac{\vec{F}_i(t)+\vec{F}_i(t+\Delta t)}{2M_i}\Delta t \end{aligned} \tag{2.15}$$

This algorithm is a time reversible symplectic transformation in the phase space, therefore total energy is well conserved in the absence of external forces [82, 84].

2.2.2 Statistical averages and size effects

According to the ergodic hypothesis the ensemble average of any dynamical quantity $A(q,p)$ is equivalent, for long times, to the time average over the trajectory of the system in phase space [82]:

$$\langle A\rangle = \int dqdp\rho(q,p)A(q,p) = \lim_{t\to\infty}\frac{1}{\tau}\int_0^\tau dtA(q(t),p(t)) \tag{2.16}$$

where the density function $\rho(q,p)$ depends on the thermodynamic ensemble considered. Within this approach molecular dynamics provides equilibrium observables, e.g. the pair correlation function $g(r)$, from time averages over atomic trajectories.

The bulk properties of a material can be calculated by simulating explicitly only a small number of atoms, of the order of 10^2, using periodic boundary conditions [78]: this amounts to consider the infinite lattice resulting from periodic replica of the simulation cell in all directions. When one atom exits from a face of the cell it enters from the opposite face, and the distance between atoms i and j is taken as the minimum distance between all the periodic replicas of i and j (minimum image convention). In this scheme physical phenomena whose correlation length exceeds the cell size can not be described properly, whereas spurious correlation between periodic replicas of the same atom can affect the accuracy of results. Moreover, investigating dynamical properties often require ensemble average that converge very slowly with simulation time, so that the size of the system plays an important role. There are even many quantities that under certain conditions are size dependent (e.g. the diffusion coefficient in the hydrodynamic regime, see section 5.2).

2.2.3 First principles molecular dynamics

In first principles molecular dynamics interatomic forces are obtained by solving the electronic structure problem during the simulation. According to Born-Oppenheimer approximation [85] the motion of slow nuclear and fast electronic degrees of freedom, having well separated typical time scales, can be conveniently decoupled:

$$M_I\ddot{\vec{R}}_I = -\nabla_{\vec{R}_I}U(\vec{R}) = -\nabla_{\vec{R}_I}\langle\psi_0(\vec{r};\vec{R})|\mathcal{H}_e|\psi_0(\vec{r};\vec{R})\rangle \tag{2.17}$$

where ψ_0 is the ground state electronic wavefunction at fixed nuclear configuration,

$$\begin{aligned}
&\mathcal{H}_e\psi(\vec{r};\vec{R}) = E_e\psi(\vec{r};\vec{R}) \\
&\qquad \mathcal{H}_e = \sum_i -\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i<j}\frac{e^2}{|\vec{r}_i-\vec{r}_j|} + \sum_{i,I}\frac{-Z_Ie^2}{|\vec{r}_i-\vec{R}_I|} + \sum_{I<J}\frac{Z_IZ_Je^2}{|\vec{R}_I-\vec{R}_J|}
\end{aligned}$$

For every instantaneous nuclear configuration, electrons obey to a Schröedinger equation at frozen nuclei, always remaining in the ground state. This amounts to neglecting possible electronic excitations due to coupling with the nuclear motion (non adiabatic effects). On the other hand nuclei move according to classical dynamics, with the total electronic energy as the interaction potential $U(\vec{R})$. The ground state electron wavefunction $\psi_0(\vec{r};\vec{R})$ depends only parametrically on $\vec{R}$, therefore the interatomic forces can be calculated by applying the Hellmann-Feynman theorem [86, 87]

$$M_I\ddot{\vec{R}}_I = -\nabla_{\vec{R}_I}\langle\psi_0|\mathcal{H}_e|\psi_0\rangle = -\langle\psi_0|\nabla_{\vec{R}_I}\mathcal{H}_e|\psi_0\rangle \tag{2.18}$$

In the Born-Oppenheimer scheme ψ_0 is computed at every step of the molecular dynamics simulation, and the atomic positions are updated according to the forces in Eq. 2.18. Ab initio molecular dynamics is computationally expensive due to the solution of the electronic problem: a convenient approach to the latter is density functional theory (DFT), which is presented in the next section. Explicit inclusion of the electronic structure is mandatory for systems where the chemical bonding between atoms is not easily described by empirical force fields. Besides molecular dynamics, the calculation of interatomic forces $\vec{F}_i$ is applied to obtain the equilibrium geometry of a system: several algorithms exist in which fictitious dynamics are exploited to minimize the energy reaching the equilibrium geometry and also to find different local minima. In the present work we employed geometry optimization techniques based on the BFGS algorithm [63] for systems up to 10^2 atoms and simulated annealing [88] for bigger systems.

2.2.4 Density functional theory

The most widespread method for the calculation of the ground state electronic structure in condensed matter is density functional theory (DFT) within the Kohn-Sham scheme [89, 90], based on the two celebrated theorems of Hohenberg and Kohn [91] and Kohn and Sham [92]. The first Hohenberg and Kohn theorem states that the ground state energy of the electronic Hamiltonian $\mathcal{H}_e$ is a functional of the electronic density as (we do not explicitly consider the spin):

$$E_0 = \langle \psi_0 | \mathcal{H}_e | \psi_0 \rangle = \int d^3\vec{r} V_{ext}(\vec{r}) n_0(\vec{r}) + F[n_0]$$

$$n_0(\vec{r}_1) = N \int d^3(\vec{r}_2) \dots d^3(\vec{r}_N) |\psi_0(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2$$

where V_{ext} is the "external" potential acting on electrons (usually the Coulomb interaction with nuclei) and $F[n_0]$ is a universal functional of the ground state electronic density, independent from V_{ext}. This formally maps the calculation of E_0 from a $3N$ variables problem (considering the coordinates of all the electrons, needed to describe ψ_0) to a 3-variables problem, at the expense of the introduction of $F[n_0]$. The second Hohenberg and Kohn theorem states that E_0 has a variational character:

$$\min_n \langle \mathcal{H}_e \rangle [n] = E_0[n_0] \tag{2.19}$$

i.e. the minimum energy E_0 corresponds to the true ground state density n_0. The Kohn-Sham theorem states that it is possible to switch from the true N-electrons Hamiltonian $\mathcal{H}_e$ to an effective one electron Hamiltonian $\mathcal{H}_{KS}$ involving just a local effective potential V_{KS}:

$$\mathcal{H}_{KS} \phi_i(\vec{r}) = \left[-\frac{\hbar^2}{2m_e} \nabla^2 + V_{KS}(\vec{r}) \right] \psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}) \tag{2.20}$$

with

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \tag{2.21}$$

The real and the fictitious system share the common total density n, which is obtained by occupying the first N one electron Kohn-Sham orbitals ψ_i

$$n(\vec{r}) = \sum_{i=1}^{N} |\psi_i(\vec{r})|^2 \tag{2.22}$$

The effective potential V_{KS} reads

$$V_{KS}(\vec{r}) = V_{ext}(\vec{r}) + V_H(\vec{r}) + V_{xc}(\vec{r}) \tag{2.23}$$

where V_H is the Hartree potential due to the mean field electron-electron interaction, and V_{xc} is the exchange-correlation potential, which includes the most complex part of the many-body electron-electron interaction:

$$\begin{aligned} V_{ext}(\vec{r}) &= \sum_I \frac{-Z_i e^2}{|\vec{r} - \vec{R}_I|} \\ V_H(\vec{r}) &= \int d^3\vec{r}' \frac{n(\vec{r}')e^2}{|\vec{r} - \vec{r}'|} \\ V_{xc}(\vec{r}) &= \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} \end{aligned} \tag{2.24}$$

where $\delta/\delta n$ is the functional derivative with respect to the density. The exchange-correlation energy functional $E_{xc}[n]$ is the unknown part of the total ground-state energy of the interacting system

$$E[n] = T_{KS}[n+] \int d^3 n(\vec{r}) V_{ext}(\vec{r}) + \frac{1}{2} \int d^3\vec{r} n(\vec{r}) V_H(\vec{r}) + E_{xc}[n] \tag{2.25}$$

T_{KS} is the kinetic energy of the auxiliary Kohn-Sham system of (formally) independent electrons

$$T_{KS}[n_0] = \sum_{i=1}^{N} -\frac{\hbar^2}{2m_e} \langle \psi_i | \nabla^2 | \psi_i \rangle \tag{2.26}$$

Thus E_{xc} includes the many-body part of both the Coulomb interaction and the kinetic energy. The Kohn-Sham theorem reduce the variational determination of the minimum of $E_0[n]$ to the solution of the single-particle equation. The difficulties of the many electrons problem are not eliminated, rather they are concentrated in the building of the unknown functional $E_{xc}[n_0]$: DFT has great effectiveness because good approximations for this functional are available, as shown in the next section. The Kohn-Sham equations are non linear, because the effective potential V_{KS} depends upon the solutions ϕ_i through the total density, thus they are usually solved by means of iterative methods until self-consistency is reached [93].

The term E_{xc} contributes just a little fraction to the total electronic energy. Nevertheless, it is fundamental because it is the largest fraction of the bonding energy which keeps together the atoms in condensed matter. The simplest approximation to E_{xc}, yet remarkably effective, is the local density approximation(LDA) [92]

$$E_{xc}^{LDA} = \int d^3\vec{r} n(\vec{r}) \epsilon_{xc}^{LDA}(n(\vec{r})) \tag{2.27}$$

where $\epsilon_{xc}^{LDA}[n_0]$ is the exchange-correlation energy of one electron in a uniform infinite jellium of density n_0, obtained from very accurate Quantum-Monte Carlo calculations [94]. The LDA satisfies many requirements on the exact E_{xc} functional, like the normalization of the exchange-correlation hole, leading to unexpectedly good description also of non-uniform systems because of extensive error-cancellation [89, 90]. The agreement with experiments can be improved in some cases including in E_{xc} also a degree of non locality through the gradient of the density (generalized gradient approximations, GGA), according to the general expression

$$E_{xc}^{GDA}[n] = \int d^3 f(n(\vec{r}), \nabla n(\vec{r})) \tag{2.28}$$

Compared to LDA, the GGA functional generally provide a less uniform density, improved cohesion and activation energies and more accurate energy difference between the crystalline phases of the material. moreover often LDA underestimates and GGA overestimates the bond lengths. The functionϵ_{xc}^{LDA} is uniquely defined, whereas various possible f functions have been proposed in the literature: in our calculations we adopted the PBE [95].

The PBE-GGA expression for the exchange-correlation energy is free of empirical parameters, and gives total energy dependent properties in good agreement with experiment [96]. The PBE-GGA functional has been extremely influential, both for performing actual calculations and as a basis for functional involving higher derivatives and exact exchange [97]. As both the density and the gradient can be constant in the homogeneus electron gas limit, there can be no unique GGA and the constraints of the PBE-GGA are not sufficient to uniquely define the functional. The PBE functional form is based on a numerical GGA [95] where a model of the exchange correlation hole was constructed to satisfy known exact hole constraints. The constraints are satisfied using a sharp real space cutoff and a dampign function, which were choices of the authors [95] and different choiches would lead to different functionals. The GGAs write the exchange energy density per particle as

$$\epsilon_x(n,s) = \epsilon_x(n) F_x(s) \tag{2.29}$$

where $\epsilon_x(n)$ is the LDA exchange energy density per particle and F_x is the enhancement factor due to density gradients and is dependent on the reduced density gradient

$$s = \frac{|\nabla n(\vec{r})|}{2(3\pi^2)^{1/3} n(\vec{r})^{4/3}} \tag{2.30}$$

The PBE enhancement factor is given as

$$F_x^{PBE}(x) = 1 + \kappa \left(1 - \frac{1}{1 + x/\kappa}\right) \tag{2.31}$$

where $x = \mu p$ and $p = s^2$. The parameters of Eq. 2.31, $\mu = 0.2195$ and $\kappa = 0.804$, are determined to ensure that the exchange gradient correction cancels for the PBE correlation as $s \to 0$ and to ensure that the local Lieb-Oxford bound is obeyed. The latter is an exact quantum-mechanical property of any Coulomb interacting system, and states that [98]

$$E_{xc}[n] \geq -C \int d^3 r n^{4/3} \tag{2.32}$$

In comparison with LDA, PBE-GGA tend to improve total energies, atomization energies, energy barriers and structural energy differences. The PBE-GGA for the exchange-correlation energy has been succesfully used in DFT simulations of phase change materials [7, 8, 26, 33]. The structural and dynamical properties of crystalline, liquid and amorphous phases are in good agreement with experimental data (see section 1.2), but several details are open to improvement. For example, known drawback of PBE-GGA and in general of all the GGA formulations, is that these functionals expand and soften atomic bonds. Akola *e*t al. [99] found that indeed bond lengths in models of amorphous GST and GeTe generated by DFT using the PBE-GGA are longer than those experimentally measured. This discrepancy can be reduced [99] by using the GGA-PBEsol [100] and the TPPS metafunctional [101].

2.2.5 The Quickstep algorithm

Density functional theory is an efficient framework for electronic structure calculations. Solving the Kohn-Sham equation (Eq. 2.20) is an algebric problem, namely the diagonalization of a matrix obtained by expanding the eigenvectors ψ_i in certain basis functions. When dealing with condensed matter, the eigenvectors ψ_i are usually expanded in plane waves; this choice allows a straighforward calculation of the Hartree potential V_H and of the exchange-correlation energy E_{xc}. On the other hand, for systems like isolated molecules the eigenvectors ψ_i are usually expressed as Gaussian functions localized on individual atoms. In this case, the number of basis functions is much smaller compared to the number of plane waves needed, so that the ortogonalization of the eigenvectors ψ_i is simpler and matrix operations become less expensive in terms of computational load. However, solving the Poisson equation for the Hartree potential $\nabla^2 V_H = -4\pi n(\vec{r})$ is less demanding by using plane waves instead of localized basis functions. One way to combine the benefits of the two approaches is the mixed basis set GPW (*hybrid Gaussian and plane waves*) [102] implemented in the Quickstep (QS) algorithm [103].

In this scheme the wavefunctions ψ_i are expanded in a Gaussian basis set, while the electronic density is expanded in terms of an auxiliary plane waves basis set as

$$n(\vec{r}) = \frac{1}{\Omega} \sum_{\vec{G}} \tilde{n}(\vec{G}) e^{i\vec{G}\cdot\vec{r}} = \sum_{\mu\nu} P^{\mu\nu} \varphi_\mu(\vec{r}) \varphi_\nu(\vec{r}) \tag{2.33}$$

where $\psi_i = C_i^\mu \varphi_\mu$ and $P^{\mu\nu} = \sum_i C_i^\mu C_i^\nu$ is the element of the density matrix $\bar{\bar{P}}$, Ω is the unit cell volume and $\vec{G}$ is a reciprocal lattice vector limited by the cutoff imposed into the plane wave expansion. Besides, also in Eq. 2.33 $\varphi_\mu(\vec{r}) = \sum_i d_{i\mu} g_i(\vec{r})$, where $g_i(\vec{r})$ are Gaussian functions with contraction coefficients $d_{i\mu}$. Once the representation of the density in a plane waves basis on a grid is known, the computational load needed for the calculation of the Hartree energy scales linearly with system size, and the Poisson equation can be solved by a Fourier transform. The latter operation can easily take advantage of the very rapid algorithms available for computing Fourier transforms (FT), like FFTW [104].

Using this dual representation, the DFT energy can be written as

$$\begin{aligned} E[n] &= E_T[n] + E_V[n] + E_H[n] + E_{XC}[n] + E_{II} \\ &= \sum_{\mu\nu} P^{\mu\nu} \langle \varphi_\mu(\vec{r})| -\frac{1}{2}\nabla^2 |\varphi_\nu(\vec{r})\rangle + \sum_{\mu\nu} P^{\mu\nu} \langle \varphi_\mu(\vec{r})| V_{PP}^{loc}(r) |\varphi_\nu(\vec{r})\rangle \\ &+ \sum_{\mu\nu} P^{\mu\nu} \langle \varphi_\mu(\vec{r})| V_{PP}^{nl}(\vec{r},\vec{r}^{\,\prime}) |\varphi_\nu(\vec{r}^{\,\prime})\rangle + 2\pi\Omega \sum_{\vec{G}} \frac{n^*(\vec{G})n(\vec{G})}{\vec{G}^2} \\ &+ \int \varepsilon_{XC}(\vec{r}) d\vec{r} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \end{aligned} \tag{2.34}$$

where $E_T[n]$ is the kinetic energy of the electrons, $E_V[n]$ is the energy due to the interaction of the electrons with the ionic cores, $E_H[n]$ is the Hartree energy, $E_{XC}[n]$ is the exchange-correlation energy and E_{II} is the energy coming from interactions between ions. Usually only valence electrons are taken into account. The interaction between ionic cores and valence electrons can be described by pseudopotentials V_{PP}, built from all-electrons calculations. In the Quickstep scheme, pseudopotentials are written as linear combination of Gaussian functions as proposed by Goedecker, Teter and Hutter [105]. Once the electronic problem has been solved, it is possible to evaluate the forces acting on the ions by explicitly calculating the gradient of the GPW energy defined in Eq. 2.34 with respect to the atomic positions.

The evaluation of the electronic ground state requires the minimization of the electronic energy with respect to the single particle orthonormal orbitals or with respect to the single particle density matrix. The solution of the Kohn-Sham equation can be also considered as a global optimization problem of the functional of the energy, which is indeed a function in a multidimensional space of the coefficients of the expansion of the ψ_i in the finite basis set. Thus, within the QS framework the Kohn-Sham problem is tackled using different techniques for the minimization of functionals in multidimensional spaces, like the direct inversion of the iterative subspace (DIIS) [106].

Beyond Born-Oppenheimer MD

In Born-Oppenheimer Molecular Dynamics (BOMD) the potential energy of the system is minimized with respect to the Kohn-Sham orbitals under the holonomic constraint

$$\langle \psi_i(\vec{r}) | \psi_j(\vec{r}) \rangle = \delta_{ij} \tag{2.35}$$

by a self consistent electronic structure calculations. This procedure has to be repeated at each step of the dynamics, so that in general BOMD is computationally expensive. An alternative approach has been proposed in 1985 by Car and Parrinello [107]. This technique is based on the transformation of the Kohn-Sham eigenstates in classical fields. These fields are associated to a fictitious kinetic energy in the Lagrangian of the system as

$$L = \sum_i \frac{1}{2}\mu|\psi_i|^2 + \sum_I \frac{1}{2} M_I \dot{\vec{R}}_I^2 - E_{DF}[n(\vec{r}), \vec{R}] \tag{2.36}$$

The Eulero-Lagrange equations obtained from Eq. 2.36 produce a dynamics in which the fields ψ_i follow adiabatically the ion motion by performing small oscillations around the Born-Oppenheimer surface defined by E_{DFT}. The adiabatic separation between electron fields and the ionic motion is ensured by choosing the fictitious mass μ sufficiently small, so that the characteristic oscillation frequency of the fields ψ_i are much larger than the characteristic vibrational frequencies of the ions. This scheme avoids the self-consistent solution of the Kohn-Sham problem at every molecular dynamics step, but there is a price to be paid. In fact, the time step used for the integration of the equation of motions is of the order of 1/10-1/100 smaller with respect to what is typically used in BOMD.

Recently, Kühne *et al.* [108] proposed a novel approach to first principles molecular dynamics that combines the advantages of both BOMD and CPMD. The electronic states are propagated without self-consistent cycles (in a CPMD -like fashion) while maintaining time steps typical of BOMD.

In this scheme, the wavefunctions are self-consistently calculated only for the first moleculary dynamics steps. Then, the density matrix $\bar{\bar{P}} = \bar{\bar{C}}\bar{\bar{C}}^T$ is propagated using the ASPC (*always stable predictor corrector*) algorithm [109]. The coefficients of the expansion of the Kohn-Sham states in the local basis $\bar{\bar{C}}$ are calculated via the following algorithm (*predictor*):

$$\bar{\bar{C}}^p(t_n) \cong \sum_{m=1}^{K} (-1)^{m+1} m \frac{\binom{2K}{K-m}}{\binom{2K-2}{K-1}} \underbrace{\bar{\bar{C}}(t_{n-m})\bar{\bar{C}}^T(t_{n-m})}_{=\bar{\bar{P}}(t_{n-m})} \bar{\bar{S}}(t_{n-m})\bar{\bar{C}}(t_{n-1}) \tag{2.37}$$

where $\bar{\bar{S}}$ is the overlap matrix. In Eq. 2.37 there is a linear combination of the matrixes $\bar{\bar{P}}$ and $\bar{\bar{S}}$ at different t_m subsequent steps. Since $\bar{\bar{P}}\bar{\bar{S}}$ is invariant with respect to unit transformations of the coefficients $\bar{\bar{C}}$, it is easier to predict then the $\bar{\bar{C}}$ themselves. Once the matrix $\bar{\bar{C}}^p(t_n)$ has been calculated, a corrector step is applied, in order to minimize the error in the propagation:

$$\bar{\bar{C}}(t_n) = \omega MIN[\bar{\bar{C}}^p(t_n)] + (1-\omega)\bar{\bar{C}}^p(t_n) \tag{2.38}$$

where

$$\omega = \frac{K}{2K-1} \tag{2.39}$$

and $MIN[\bar{\bar{C}}^p(t_n)]$ is a single minimization step. This procedure inevitably introduces an error into the calculation of the ground state density, since $\bar{\bar{C}}(t_n)$ is only an approximation eigenfunction of $H[n^p]$ into the subspace sampled by the finite basis set. Thus, the error in the evaluation of the forces acting on the ions is negligible only if $\bar{\bar{C}}(t_n)$ is always sufficiently close to the ground state. However, it turns out that this error can be approximated to a white noise.

The precision by which this dynamics follows the Born-Oppenheimer surface depends on the system itself, but in any case the dynamics is dissipative, as

$$\vec{F}_{PC} = \vec{F}_{BO} - \gamma_D \dot{\vec{R}}_I \tag{2.40}$$

where $\vec{F}_{PC}$ are the forces generated by the dynamics described in here, $\vec{F}_{BO}$ are the forces coming from a pure Born-Oppenheimer dynamics and γ_D is a friction coefficient. In Ref. 108 it has been shown how to compensate for this frictional term by using a thermostat on the ions thus recovering a canonical sampling of atomic velocities. The method proposed by Kühne and Parrinello [108] has been implemented in the CP2K package [110].

2.2.6 Classical molecular dynamics

In classical molecular dynamics, the electronic problem is not, at least explicitly, taken into account. Thanks to the Born-Oppenheimer approximation, ionic and electronic degrees of freedom are decoupled, and one can assume a functional form for the interatomic potential. For example, an expression of the interatomic potential $U(\vec{r})$ analytical in the atomic positions and dependent on a number of parameters can be fitted e.g. to an experimental or theoretical available dataset. Since the potential energy is analytically known, evaluation of forces, which usually is the most time consuming part of an MD simulation, is straightforward, order of magnitude faster than solving Schroedinger equation for the electronic system at each or every several steps like in first principles MD. Of course, while the latter relies on solid physical bases, in classical MD the whole reliability of the simulation depends on the choice of the interatomic potential.

2.3 Thermodynamic Integration of the free energy

To validate our NN potential for GeTe we computed the melting temperature T_m. A proper calculation of $T_m$requires comparing the free energies of the liquid and the solid phases. The method of choice in this context is known as thermodynamic integration [82]. Once a coexistence point (in the PT space) is located, it is possible to obtain the melting line by the so called Gibbs-Duhem integration method of Kofke [111].

The second law of thermodynamics states that for a closed system with energy E, volume V, and number of particles N, the entropy S is at maximum when the system is in equilibrium. That means that in the NVT ensemble, the Helmoltz free energy $F = E - TS$ is at a minimum in equilibrium. If for example we wish to know which of the two phases (solid or liquid) is stable at a given temperature and density, we have to compare the Helmoltz free energies F_s and F_l of the solid and the liquid. Entropy, free energy and relative quantities are not simply averages of functions of the phase space coordinates of the system. To compute the free energy of a system at given temperature and density, we should find a reversible path in the V-T plane that links the state under consideration to a state whose free energy is known. The change in F along the path can then simply be evaluated by integration of the equations

$$\begin{aligned} \left(\frac{\partial F}{\partial V}\right)_{NT} &= -P \\ \left(\frac{\partial F/T}{\partial 1/T}\right)_{NV} &= E \end{aligned} \tag{2.41}$$

The formalism that allows calculating the free energy difference between a reference state whose free energy is analytically known and the real phase is called Krikwood's coupling parameter method [112]. There are only very few thermodynamic states for which the free energy of a substance is known. We chose as reference systems the Einstein crystal for the solid phase, and the Lennard-Jones liquid for the liquid phase. Let us consider a N-particle system with a potential energy function U. We assume that U depends linearly on a coupling parameter λ such that, for $\lambda = 0$, U corresponds to the potential energy of our reference system U_{ref} while for $\lambda = 1$ we recover the potential energy of the system of interest, that in our case is described by the NN potential and so we indicate it with U_{NN}. Thus

$$U_\lambda = \lambda U_{NN} + (1 - \lambda) U_{ref}. \tag{2.42}$$

We now write the partition function for a system with a potential energy function that correspond to a value of λ between 0 and 1 as

$$Q_{N,V,T,\lambda} = \frac{1}{\xi^{3N} N!} \int e^{-\beta U_\lambda} d\vec{r}^N \tag{2.43}$$

where ξ is the thermal de Broglie wavelength and $\beta = \frac{1}{K_B T}$. The derivative of the Helmoltz free energy F_λ with respect to λ can the be written as an ensemble average

$$\left(\frac{\partial F_\lambda}{\partial \lambda}\right)_{NVT} = \left\langle \frac{\partial U_\lambda}{\partial \lambda} \right\rangle \tag{2.44}$$

where $\langle \ldots \rangle_\lambda$ denotes an ensemble average for a system with a potential energy function

U_λ defined in Eq. 2.42. If we integrate Eq. 2.44 we obtain an expression for the free energy difference between the reference state and the actual state

$$F_{NN} = F_{ref} + \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial U_\lambda}{\partial \lambda} \right\rangle d\lambda \tag{2.45}$$

It should be noticed that Eq. 2.45 expresses a free energy difference in terms of an ensemble average, which, unlike a free energy, can be calculated directly in a simulation. In particular, we obtained the derivative of the energy in a series of equilibrium MD simulations.

2.3.1 The solid phase

As the reference system for the crystalline phase of GeTe we chose an Einstein crystal, i.e. a solid of non-interacting particles that are all coupled to their respective lattice sites by harmonic springs. In thermodynamic integration we switch on these spring constants and consequently switch off the intermolecular interactions by varying the λ parameter that mixes the NN interatomic potential U_{NN} and the Einstein crystal potential U_{ref}. Because GeTe is a binary system, we need to assign to Ge and Te atoms different coupling constants. The Einstein crystal potential (or internal energy) is

$$U_{ref} = U(\vec{r}_0^N) + \frac{\alpha_{Ge}}{2}\sum_{i=1}^{N_{Ge}}(\vec{r}_i - \vec{r}_{0,i})^2 + \frac{\alpha_{Te}}{2}\sum_{i=1}^{N_{Te}}(\vec{r}_i - \vec{r}_{0,i})^2 \tag{2.46}$$

where $\vec{r}_0^N$ is the set of coordinates of the minimum energy configuration (the equilibrium atomic coordinates) and $U(\vec{r}_0^N)$ its potential energy, that can be evaluated by averaging in a MD run the potential energy of the NN system in the minimum of the equation of state at the temperature of interest. The second term is the harmonic potential of the N oscillators, where α_{Ge} and α_{Te} are the associated force constants These force constants can be adjusted to optimize the accuracy of the numerical integration of Eq. 2.45. The integration is optimal if the interactions in the pure Einstein crystal differ as little as possible from those in the actual GeTe crystal. This suggest that α_{Ge} and α_{Te} should be chosen such that the mean-squared displacement for $\lambda = 0$ and $\lambda = 1$ are equal. For a non interacting Einstein crystal, the mean square displacement is given by

$$\langle r_k^2 \rangle_{\lambda=0} = \frac{3}{\beta \alpha_k} \tag{2.47}$$

so the following condition holds for α_k, where k can be Ge or Te:

$$\alpha_k = \frac{3}{\beta \langle \frac{1}{N} \sum_{i=1} N_k (\vec{r}_{i,k} - \vec{r}_{i,0,k})^2 \rangle_{\lambda=1}} \tag{2.48}$$

where $\vec{r}_{i,0,k}$ is the equilibrium lattice position of atom i. We obtained $\alpha_{Ge} = 9.555$ N/m and $\alpha_{Te} = 12.879$ N/m. The potential energy difference between the real and the reference crystal can be calculated along a reversible pathway that links the two systems by changing λ. It should be notice that the path must not cross a first order phase transition. The potential U_λ will then read

$$\begin{aligned}
U_\lambda &= \lambda U_{NN} + (1-\lambda)U_{ref} \\
&= \lambda U_{NN} + (1-\lambda)[U(\vec{r}_0^N) + \frac{\alpha_{Ge}}{2}\sum_{i=1}^{N_{Ge}}(\vec{r}_i - \vec{r}_{0,i})^2 + \\
&\frac{\alpha_{Te}}{2}\sum_{i=1}^{N_{Te}}(\vec{r}_i - \vec{r}_{0,i})^2]
\end{aligned}$$

where for $\lambda = 0$ we have the pure Einstein crystal while for $\lambda = 1$ we recover the actual GeTe crystal. In the Einstein solid, the equilibrium lattice positions are fixed to an absolute frame, so that the energy is not invariant with respect to a translation of the crystal as a whole. On the other hand, when $\lambda = 1$ the center of mass of the system drifts due to inaccuracies in the integration of the equation of motion, and the particles may be allowed to move far away from their absolute equilibrium lattice positions, so that U_{ref} becomes uncontrollably large. We thus need to perform simulations under the constraint that the center of mass of the solid is fixed. This constraint introduces two additional terms in Eq. 2.45. The first accounts for the center of mass constraint, so that the Helmoltz free energy of an Einstein crystal under the above mentioned constraint can be written as as [113]

$$\begin{aligned}
\frac{\beta F_{ref}}{N_{GeTe}} &= 3\ln\xi_{Ge} + 3\ln\xi_{Te} - \frac{3N_{Ge}}{2N_{GeTe}}\ln\left(\frac{2\pi}{\beta\alpha_{Ge}}\right) - \frac{3N_{Te}}{2N_{GeTe}}\ln\left(\frac{2\pi}{\beta\alpha_{Te}}\right) \\
&- \frac{3}{2N_{GeTe}}\ln\left(\frac{\beta\alpha_{Ge}}{2\pi N_{Ge}\mu_{Ge}^2}\right) - \frac{3}{2N_{GeTe}}\ln\left(\frac{\beta\alpha_{Te}}{2\pi N_{Te}\mu_{Te}^2}\right) \\
&- \frac{3}{2N_{GeTe}}\ln\left(\frac{\beta h^2}{2\pi(N_{Ge}m_{Ge} + N_{Te}m_{Te})}\right)
\end{aligned} \tag{2.49}$$

where μ_i is the fractional mass of each species, ξ_i is the thermal de Broglie wavelength of the i-th species given by $\xi_i = h/\sqrt{2\pi m_i k_B T}$ and N_{GeTe} is the number of GeTe formula units. The other additional term originates from the finite system-size corrections associated with the center of mass constraint, and it is [113]

$$\begin{aligned}
\frac{\beta F_{FS}}{N_{GeTe}} &= \frac{1}{N_{GeTe}}\ln\left(\frac{N_{GeTe}}{V}\right) + \\
&\frac{3}{2N_{GeTe}}\ln\left(\frac{\beta h^2}{2\pi(N_{Ge}m_{Ge} + N_{Te}m_{Te})}\right)
\end{aligned} \tag{2.50}$$

We now rewrite the final expression for the Helmoltz free energy of crystalline GeTe in β units:

$$\frac{\beta F_{NN}}{N_{GeTe}} = \frac{\beta F_{ref}}{N_{GeTe}} + \frac{\beta}{N_{GeTe}} \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial U_{\lambda}}{\partial \lambda} \right\rangle d\lambda + \frac{\beta F_{FS}}{N_{GeTe}} \tag{2.51}$$

with F_{ref} and F_{FS} given by Eq. 2.50 and Eq. 2.51 respectively. In order to compute the integral in Eq. 2.51, we performed ten simulations for ten different values of λ chosen in order to allow a ten point computation of the integral by Gauss-Legendre quadrature, an integration scheme that accurately approximates the value of a definite integral of the non-singular function $f(x)$ as

$$\int_{x1}^{x2} f(x)dx = \sum_{j=1}^{N} w_j f(x_j) \tag{2.52}$$

where the abscissas x_j for the quadrature of order N are given by the roots of the Legendre polynomials $P_N(x)$ symmetric around zero and weights w_j are given by

$$w_j = \frac{2}{(1 - x_j^2)[P'_N(x_j)]} \tag{2.53}$$

2.3.2 The liquid phase

In order to compute the Helmoltz free energy of liquid GeTe, we proceed as discussed in the case of the crystal, but for the choice of another reference system. We chose the Lennard-Jones (LJ) liquid, whose free energy at various temperatures and densities were computed using the equation of state (EOS) proposed by Johnson *et al.* [114]. These authors determined semi-empirical parameters for the excess Helmoltz free energy F_{xs} with respect to the ideal gas of the cut and shifted LJ interatomic potential given by

$$U_{LJ}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r_c}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} + \left(\frac{\sigma}{r_c}\right)^{6} \right] & r \leq r_c \\ 0 & r \geq r_c \end{cases} \tag{2.54}$$

The EOS of Johnson *et al.* allows evaluating $F_{ref,LJ}$ as:

$$F_{ref,LJ} = F_{id} + F_{xs} + F_{corr} + F_{size} \tag{2.55}$$

where F_{id} is the ideal gas mixture Helmoltz free energy

$$\begin{aligned} \frac{\beta F_{id}}{N} = {} & X_{Ge}[3\ln(\xi_{Ge}) + \ln(N/V) - 1] + \\ & X_{Te}[3\ln(\xi_{Te}) + \ln(N/V) - 1] + [X_{Ge}\ln(\xi_{Ge}) + X_{Te}\ln(\xi_{Te})] \end{aligned} \tag{2.56}$$

where $X_i = \frac{N_{Ge/Te}}{N}$ is the atomic fraction of each species. F_{xs} is the excess Helmoltz free energy, given by [114]

$$\frac{\beta F_{xs}}{N} = \sum_{i=1}^{8} \frac{a_i \rho_i^{LJ}}{i} + \sum_{i=1}^{6} b_i G_i \tag{2.57}$$

where ρ^{LJ} is the density in reduced LJ units ($\rho^{LJ} = \rho\sigma^3$ and $T^{LJ} = k_B T/\epsilon$). This is the so called modified Benedict-Webb-Rubin (MBWR) equation of state [114], that through the coefficients a_i and b_i contains 32 parameters, that depend on temperature and have been obtained by an extensive and accurate fitting of a huge set of reference data along a wide (ρ, T) region of the LJ liquid. The functions G_i contain instead exponential of the density and a single nonlinear parameter. The full form and value of all the parameters and functions contained in Eq. 2.57 can be found in Ref. 114. Using the cut and shifted form of the LJ potential of Eq. 2.54 introduces a correction to the excess free energy of Eq. 2.57, analytically computed by Johnson *et al.* [114]:

$$\frac{\beta F_{corr}}{N} = -\frac{32N\sigma^3\epsilon\beta}{9\pi V}\left[\left(\frac{\sigma}{r_c}\right)^9 - \frac{3}{2}\left(\frac{\sigma}{r_c}\right)^3\right] \tag{2.58}$$

Finally, a term accounting on finite size effects must be taken into account:

$$\frac{\beta F_{size}}{N} = (\ln 2\pi N)\frac{1}{2N} \tag{2.59}$$

We used the same set of LJ parameters σ and ϵ for all the interatomic interactions. The LJ σ parameter was determined by matching the position of the first peak of the radial distribution functions (g(r)) of the NN liquid GeTe and the LJ liquid, ensuring optimal similarity between the structure of the two liquids. The LJ ϵ parameter was chosen such that, at the selected temperature, the LJ liquid was above the critical temperature in order to avoid phase transitions which would preclude the use of thermodynamic integration. On the other hand, the liquid should not be too far from the critical temperature: in fact, the radial distribution function g(r) given by the NN potential for liquid GeTe has pronounced secondary peaks beyond the first coordination shell. Thus, a rather structured LJ liquid had to be preferred for convergence of the thermodynamic integration. The requirements are matched by choosing a LJ liquid in the proximity of the critical temperature. In our case, we chose $\rho^{LJ} = 0.75$ and $T^{LJ} = 1.3$, which assign σ and ϵ of the LJ potential. Thermodynamic integration was performed for a given reference temperature T' corresponding to the experimental T_m at normal pressure (998 K for GeTe [21]) at the experimental density of liquid GeTe at T_m (0.0339 atoms/Å^3 [21]).

2.3.3 Chemical potentials

To estimate the melting temperature we have to compare the chemical potential of the two phases. To compute the chemical potentials at constant pressure and temperature from the Helmoltz free energy computed at constant density and temperature we proceeded as follows. We computed the equation of state $P = P(\rho, T')$ for each phase at the reference temperature T'. The $P(\rho)$ isotherm was fitted to a quadratic polynomial

$$P(\rho, T') = a(T') + b(T')\rho + c(T')\rho^2 \tag{2.60}$$

Once determined the three parameters a, b and c, we can solve Eq. 2.60 with respect to the density to obtain $\rho = \rho(P, T')$. We now write the chemical potential as

$$N\mu = F + PV \tag{2.61}$$

and

$$\begin{aligned} \mu(\rho) &= \mu(\rho') + \Delta\mu \\ &= \mu(\rho') - \frac{1}{N}\left[\int_{\rho}^{\rho'} \frac{\partial F}{\partial \rho} d\rho + \frac{P(\rho)}{\rho} - \frac{P(\rho')}{\rho'}\right] \end{aligned} \tag{2.62}$$

where ρ' is the number density at the initial state point. Working out the integral in Eq. 2.62 with

$$\frac{\partial F}{\partial \rho} = \frac{P}{\rho^2} \tag{2.63}$$

and the aid of Eq. 2.60 we got

$$\mu(\rho) = \frac{F'}{N} + \left(\frac{a}{\rho'} b \ln \frac{\rho}{\rho'} + b + c(2\rho - \rho')\right) \tag{2.64}$$

where F' is the Helmoltz free energy of either liquid or solid GeTe at the reference temperature $T' = 998K$. From the equation of state $\rho = \rho(P, T)$ we finally obtain the chemical potential $\mu = \mu(P, T)$. By equating $\mu_{liq} = \mu_{liq}(P, T')$ and $\mu_{sol} = \mu_{sol}(P, T')$ we obtain the transition pressure at T'.

In order to obtain the full melting curve $T_m = T_m(P)$, one can integrate the Clausius-Clapeyron equation, for example using the numerical technique proposed be Kofke [111]. However, since the transition pressure P' is not too far from ambient pressure it is possible to compute the melting temperature at normal conditions from the slope of the Clausius-Clapeyron equation at T', i.e.

$$T_m = T' + \frac{dT}{dP}(P - P') \tag{2.65}$$

where

$$\frac{dT}{dP} = \frac{T\Delta V}{\Delta S} \tag{2.66}$$

at T', where ΔV and ΔS are the changes of volume and entropy across the phase change at the point (T', P'). Moreover $T\Delta S = \Delta E + P\Delta V$ which allows computing $\frac{dT}{dP}$ from ΔE and ΔV obtained from the simulations at the equilibrium for the liquid and the solid phase.

2.3.4 Van der Waals correction

Since the NN potential has been fitted on a database DFT-PBE, it does not include vdW interactions. Actually, the equation of state of liquid GeTe does not even show a minimum if long range vdW interactions are not taken into account in DFT-PBE calculations. Thus, in order to reproduce the equilibrium density of the liquid at T_m, an empirical Van der Waals (vdW) correction had to be added to the NN potential. We chose the vdW correction according to the scheme proposed by Grimme [115]. This is a semiempirical correction that can be applied *a posteriori* to the energy (or to the pressure) in the form

$$E_{disp} = -s_6 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{C_6^{ij}}{R_{ij}^6} \cdot \frac{1}{1 + e^{-d[(R_{ij}/R_{vdW})-1]}} \tag{2.67}$$

where C_6^{ij} is the dispersion coefficient for atoms pair ij, $s6$ is a global scaling factor, and R_{ij} is the interatomic distance. A damping function which turns off the vdW interaction at short distances is given by the term $\frac{1}{1+e^{-d[(R_{ij}/R_{vdW})-1]}}$ in Eq. 2.67 where R_{vdW} is the sum of atomic vdW radii. The values of C_6^{ij} have been obtained for a large number of elements including Ge and Te and can be found in Ref. 115. We tuned the $s6$ parameter to 0.55 to reproduce the experimental equilibrium volume of the liquid at T_m [116]. We have verified that by changing the value of $s6$ from 0.4 to 0.7 the value of the melting temperature of GeTe at ambient pressure changes by at most 5 K. The experimental equilibrium volume of the amorphous and crystalline phases are instead well reproduced by the NN potential without the need of the vdW interaction. The inability of the NN potential (and of PBE calculations as well) in reproducing the equilibrium volume of the liquid can be traced back to the presence of nanovoids in the melt [8]. In the liquid the nanovoids can coalesce and increase in size by decreasing the density which results into a reduced tensile stress upon expansion. This effect is hindered by vdW interactions. Nanovoids are also present in the amorphous phase[8], but their distribution cannot change by scaling the volume at fixed temperature because of the low atomic mobility in the amorphous phase. By investigating the properties of supercooled liquid GeTe, we verified that the distribution of voids in the liquid increases by decreasing the density. The volume occupied by the voids is computed according to the definition of Ref. 117 and the algorithm of Ref. 118. The volume fraction Ξl occupied by voids as a function of density is reported in Fig. 2.7. Clearly the empty space in the system increases as density decreases. By decreasing density small nanocavities collapse forming larger voids. The distributions of the volume of nanocavities at two different densities are shown in Fig. 2.8. At low density large voids 500 Å^3 in size are formed. A graphic representation of the liquid GeTe voids is shown in Fig. 2.9.

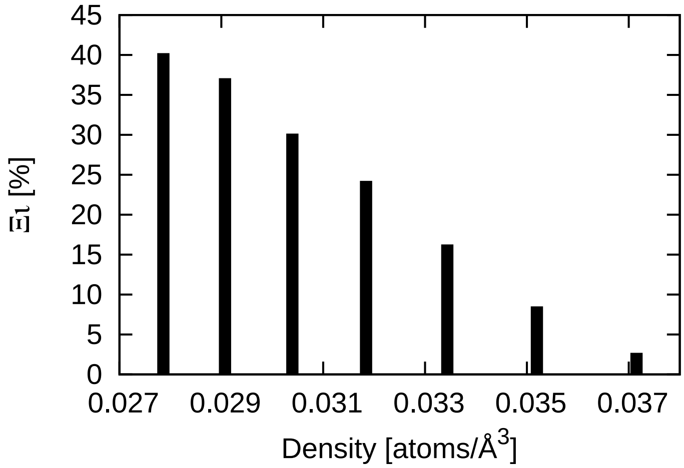

Figure 2.7: Volume fraction of liquid GeTe at 998 K occupied by voids as a function of density.

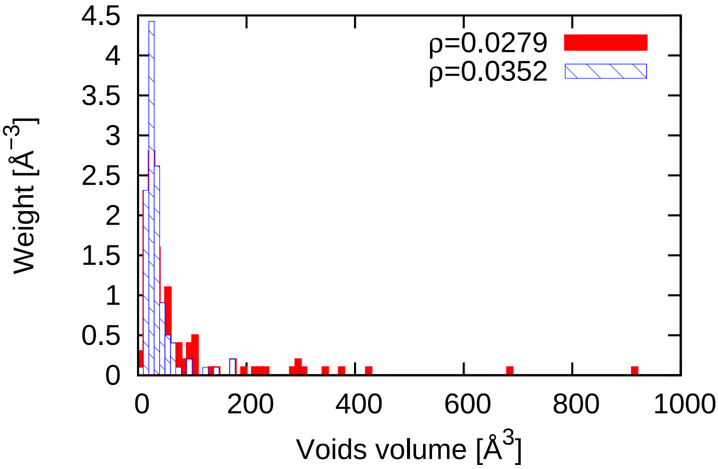

Figure 2.8: Distribution of the volume of cavities in liquid GeTe at 998 K for two different densities.

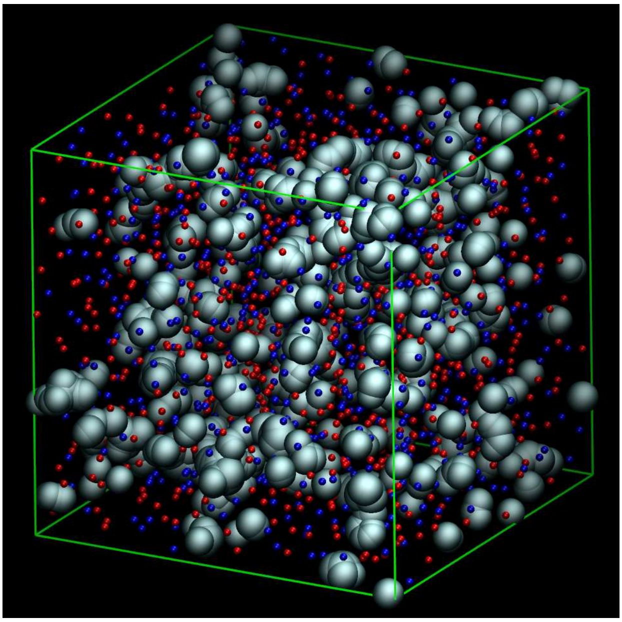

Figure 2.9: Snapshot of liquid GeTe at 1000 K at its theoretical equilibrium density (see section 3.2.2). White spheres and their superposition mark the region of space occupied by nanocavities (or voids).

The vdW correction is used in this work only in the calculation of the equations of state to assign an equilibrium volume t different temperatures, but not in the MD simulations due to uncertainties in the reliability of the correction at short interatomic distance.

Chapter 3

The Neural Network Potential for GeTe

3.1 A Neural Network potential for GeTe

To generate the NN potential, we fitted the total energies of about 30.000 configurations of 64-, 96- and 216-atom supercells computed within Density Functional Theory (DFT). We started with a relatively small dataset of about 5.000 structures, where we considered crystalline configurations, snapshots of the liquid phase and of the amorphous phase generated by quenching from the melt at ambient conditions and at different pressures up to 50 GPa. We also considered mixed crystalline/amorphous models generated by partially crystallizing the amorphous phase by means of the metadynamics technique [119]. All these configurations were generated within DFT molecular dynamics simulations at different temperatures (up to 3000 K) with the code CP2k [103] and the Perdew-Burke-Ernzerhof (PBE) [95] exchange-correlation functional. The Kohn-Sham orbitals were expanded in a Triple-Zeta-Valence plus Polarization (TZVP) Gaussian-type basis set and the charge density was expanded in a planewave basis set with a cut-off of 100 Ry to efficiently solve the Poisson equation within the Quickstep scheme [103]. Goedecker-type pseudopotentials [105] with four and six valence electrons were used for Ge and Te, respectively. Brillouin Zone (BZ) integration was restricted to the supercell Γ point. A time step of 2 fs was used for the simulations. The same scheme was applied in previous works on GeTe, GST and other phase change materials [33, 120] and previously validated by comparison with plane waves molecular dynamics simulations with the CPMD code [121]. This framework was used to generate the atomic configurations for the fitting of the NN potential and to generate the DFT models of liquid and amorphous GeTe to check to transferability of the NN potential. However, since the NN fitting relies only on tiny differences in the total energy of supercells also including configurations with different number of atoms, a very high accuracy in the total energy is mandatory. An accurate integration of the BZ is indeed needed to achieve the required convergence in the total energy for the NN fitting. To generate the energies dataset for the fitting, we thus used the atomic configurations generated with the CP2k code and computed their total energy by performing BZ integration over a dense $4 \times 4 \times 4$ Monkhorst-Pack (MP) [122] mesh for the 64-atom cell and employing meshes of a corresponding k-point density for the larger systems. To this aim we used the QUANTUM-ESPRESSO package [123]. Norm conserving pseudopotentials were employed, considering only the outermost s and p electrons in the valence shell. The Kohn-Sham orbitals were expanded in a plane waves basis up to a kinetic energy cutoff of 40 Ry. These settings ensure convergence of the total energy to 2 meV/atom. The fact that we used two different DFT setups to generate the atomic configurations and to compute their total energy for the NN fitting is immaterial as we could have generated the atomic configurations by any means including empirical interatomic potentials provided that the database is large enough to include all the relevant configurations.

This first dataset was then expanded by adding randomly distorted structures of the initial dataset at slightly different pressures and temperatures and models with slight deviations from the perfect stoichiometry. The refinement of the potential was achieved by inserting in the data set the ab initio energy of configurations generated by Molecular Dynamics simulations (see below) using the not yet refined NN potential.

The best NN fit we found employs three hidden layers with 20 nodes each. The total number of weights of our NN architecture is 8080. For each species, we have 159 symmetry functions in the input layer. From the input layer to the first hidden layer, which contains 20 nodes, we need $159 \cdot 20 = 3180$ weights. From the first to the second hidden layer we need $20 \cdot 20 = 400$ weights, and the same holds for the path from the second to the third hidden layer (from there to the output layer we just perform a summation, so there is no weight involved). So far, we have 3980 weights for species. We now add the bias weights; for each hidden layer of n nodes, we choose to add n bias weights in order to adjust the offset of the output of each node in each hidden layer. Since we have three hidden layers of 20 nodes each, we need 60 bias weights. Thus, for each species we employed 4040 weights, so that our NN potential is an analytical expression of 8080 parameters. Sigmoid activation functions were used in the nodes of the hidden layers, while a linear function was used for the output node. The dependence of the results on these choices is discussed in the next section. The local environment of each atom is defined by the value of 159 symmetry functions (see Ref. [124] for details) defined in terms of the positions of all neighbors within a distance cutoff of 6.88 Å. We checked that by decreasing the cutoff from 6.88Å to 6.00Å the structural properties of the liquid and amorphous phases do not change by inspection on the partial pair correlation functions. However, the larger cutoff turned out to be necessary to reproduce the DFT diffusion coefficient of the liquid.

The generation of the NN potential and the calculation of the forces for the MD simulations were performed with the NN code RuNNer [125]. We used the DL_POLY [126] code as MD driver. The time step for the MD runs was set to 0.2 fs, and constant temperature simulations were performed using the Berendsen thermostat [127].

3.1.1 Details on the Neural Network architecture

In order to construct our NNP, we expanded the initial input dataset in an iterative fashion, trying to sample the whole configurational space of interest for bulk GeTe. The final dataset is depicted in Fig. 3.1 as an energy versus density plot. The experimental densities of liquid, amorphous and crystalline GeTe are 0.0340 atoms/Å^3 [116] at the melting temperature T_m = 998 K, 0.0371 atoms/Å^3 in its trigonal phase stable at zero temperature [128] and 0.03327 atoms/Å^3 at zero temperature[129], respectively. The dataset is much more dense in the density region of interest for the applications. Still, it is important to sample the high density region as well, because the NN potential has to learn how to behave when bonds became shorter i.e. it must properly describe repulsive interaction at short distances. Moreover, high density configurations are useful if one would like to study the behavior of GeTe under pressure. The equilibrium configurations of all the bulk phases of GeTe lie in the lowest energy region of Fig. 3.1, below -164 eV/atom. However, in order to sample exotic configurations that indeed may be present even in ordinary MD simulations like for example liquid GeTe at very high temperature, where atoms move almost gas-like and can really explore a vast region of configurational space, it is necessary to include a huge number of high energy structures, which are not needed to be described with accuracy but contribute to establish the reliability of the NN potential.

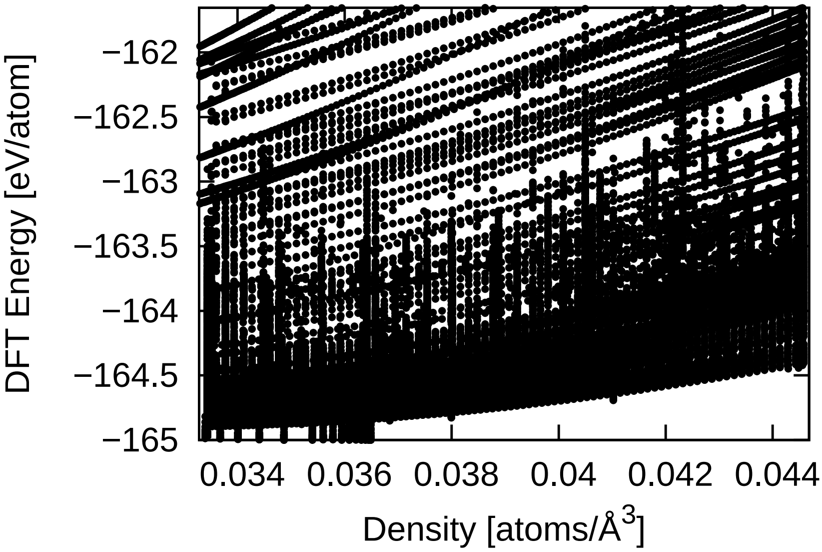

Figure 3.1: Input dataset points. DFT Energies have been calculated according to details given in section 3.1.

In Fig. 3.2 a sort of density of states for both energy and density of the input dataset is shown in panel a) and b) respectively. In constructing the dataset, it is important not to let holes in the energy-density space, as the NN can encounter numerical problems in fitting accurately two very different regions of the configurational space.

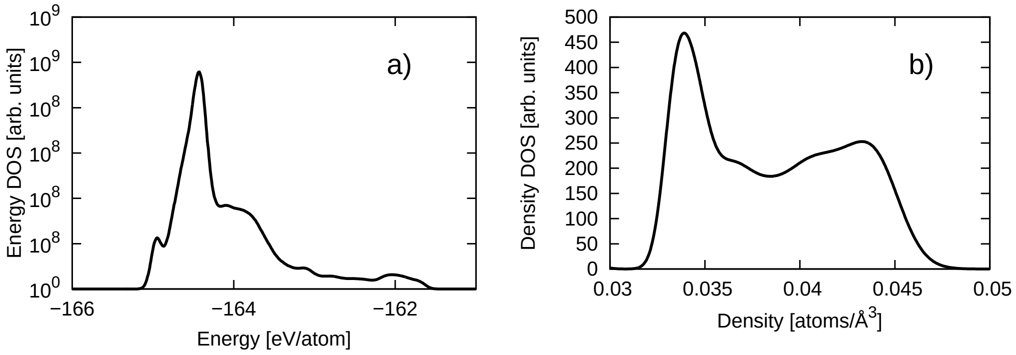

Figure 3.2: Energy a) and density b) DOS of the input dataset points.

Concerning the quality of the NN fit, there is a number of variables in the architecture of the NN that have to be carefully tuned. We shall review the most important ones with the aid of Fig. 3.3 and Fig. 3.4, where we report the accuracy of different NN fits (measured by the root means standard error RMSE) as a function of different NN variables. We underline that numerical accuracy is not sufficient to guarantee the reliability of the NN interatomic potential. The real benchmark is given instead by the comparison between physical properties of the systems obtained by the NN potential and the reference potential (in our case, the DFT data). In Fig. 3.3 and Fig. 3.4 results for the same dataset are compared with the same architecture of NN potential but for a single variable which is changed at once.

The order by which the dataset points are feed to the NN affects the quality of the fit, because the optimization algorithm used to fit the weights (the Kalman filter, see section 2.1.7) is based on online learning, that is, the NN weights are optimized after each individual training point has been taken into account. In panels a) and b) of Fig. 3.3 we compare the RMSEs of two fits, one in which the training points are fed in blocks of similar datapoint in a pseudo-random fashion, while the other has been obtained by shuffling the whole dataset in a real random fashion. The difference in the quality of the fits is evident, and similar in energies and forces. At this point we should point out that energies and forces often do not show a similar behavior with respect to a change of NN variables.

In principle, the NN weights can assume any value in any numerical range, but it is convenient to rescale them in a certain interval, in order to make the activation function application more effective. In panels c) and d) of Fig.3.3 the accuracy of different fits is compared with respect to different ranges in which the initial NN weights have been initialized. It appears that reducing the range gives better results, but in practice one has to tune the initial weights range with non linear region width of the activation function. Our best fit has been obtained with a $[-1, +1]$ range for all the weights.

Different activation functions (see section 2.1.3) can be used. In panel e) and f) of Fig.3.3 we compared different fits, each one employing a different activation function for all the nodes and layers of the NN (with the exception of the output layer, in which a linear function has been always used in this work). Because is much more difficult to reduce the RMSE of the forces than the RMSE of the energy, we have chosen the sigmoid (sigma) activation function. The smoothness and the position of the inflession point of the sigmoid function can also be tuned. In panels h) and g) of Fig.3.3 we shown how the numerical quality of the fits is affected by different choices of σ, a parameter that modifies the sigmoidals function as

$$f(x) = \frac{1}{1 + e^{-\sigma x}} + \beta \tag{3.1}$$

where β shifts the sigmoid into a different range. Our best fit has been obtained with $\sigma = 0.6$ and $\beta = +0.5$. It can be noticed that in this case energies and forces display opposite behavior with respect the choice of the smoothing. We have found that the more flexible the NN is, the better the energy RMSE gets but forces RMSE simply rises up. This behavior is not general, as each system and each NN potential constitute an unique situation that must be handled in a different way. Still, in the case of GeTe this is particularly true if we take a look to panels a) and b) of Fig. 3.4, that clearly show what happens if we increase the number of hidden layers. In fact, increasing NN flexibility results in a better RMSE for energy but a worse RMSE for forces.

Something similar happens if we increasing the number of nodes in the hidden layers, as shown in panels c) and d) of Fig. 3.4, even if in this case only the RMSE of the forces displays a clear trend. Our best fit was obtained using three hidden layers of twenty nodes each, which is a compromise in terms of flexibility.

A role of paramount importance not only in the numerical quality of the fit, but also in the transferability of the NN potential, is played by the symmetry functions (SF) choice. In panels e) and f) of Fig. 3.4 we compare results obtained with different kind of SF, together with the effect of the cutoff radius choice in panels g) and h) of Fig. 3.4. It turns out that even varying only a couple of symmetry functions of the whole (typically huge, we used 159 SF for each atomic species) set of SF can lead to important differences in the physical properties of GeTe predicted by the NN potential.

As a proof of the numerical accuracy of our best fit, we report in Fig. 3.5 the reference DFT energies versus the predicted NN energy) for the whole dataset (training and test points as well). The plot shows a remarkable agreement.

The results of the fitting process of the NN potential are summarized in Fig. 3.6. The root mean square error (RMSE) for the energy is 5.01 and 5.60 meV/atom for the training and the test set, respectively, while the RMSEs of the forces are 0.46 and 0.47 eV/Å, for the two sets. Among all structures considered, only a negligible fraction shows noticeable absolute errors, up to 25.8 meV/atom and 11.2 eV/Å for energies and forces (see Fig.

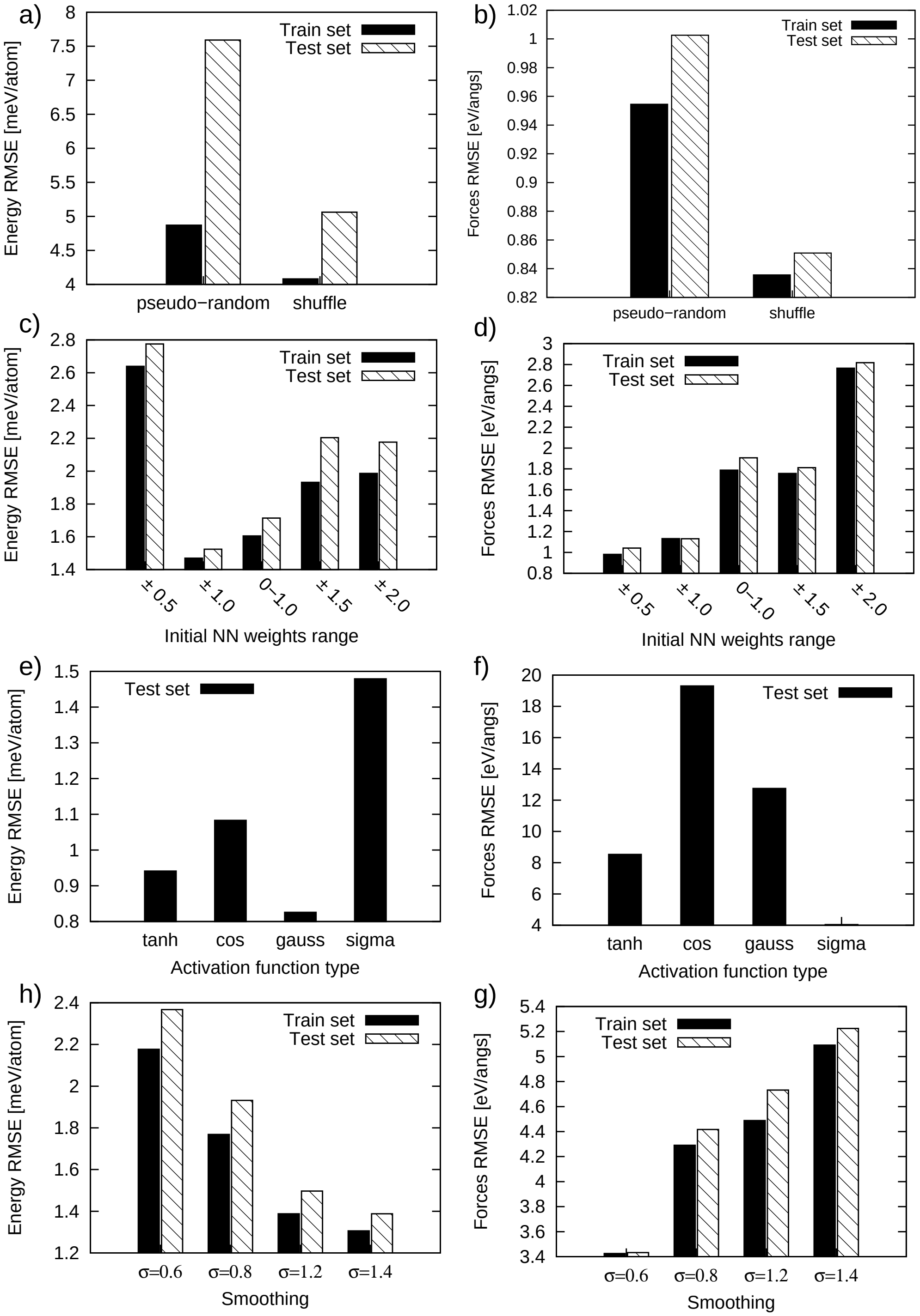

Figure 3.3: Comparison of energy and forces RMSEs for different NN architectures and parameters.

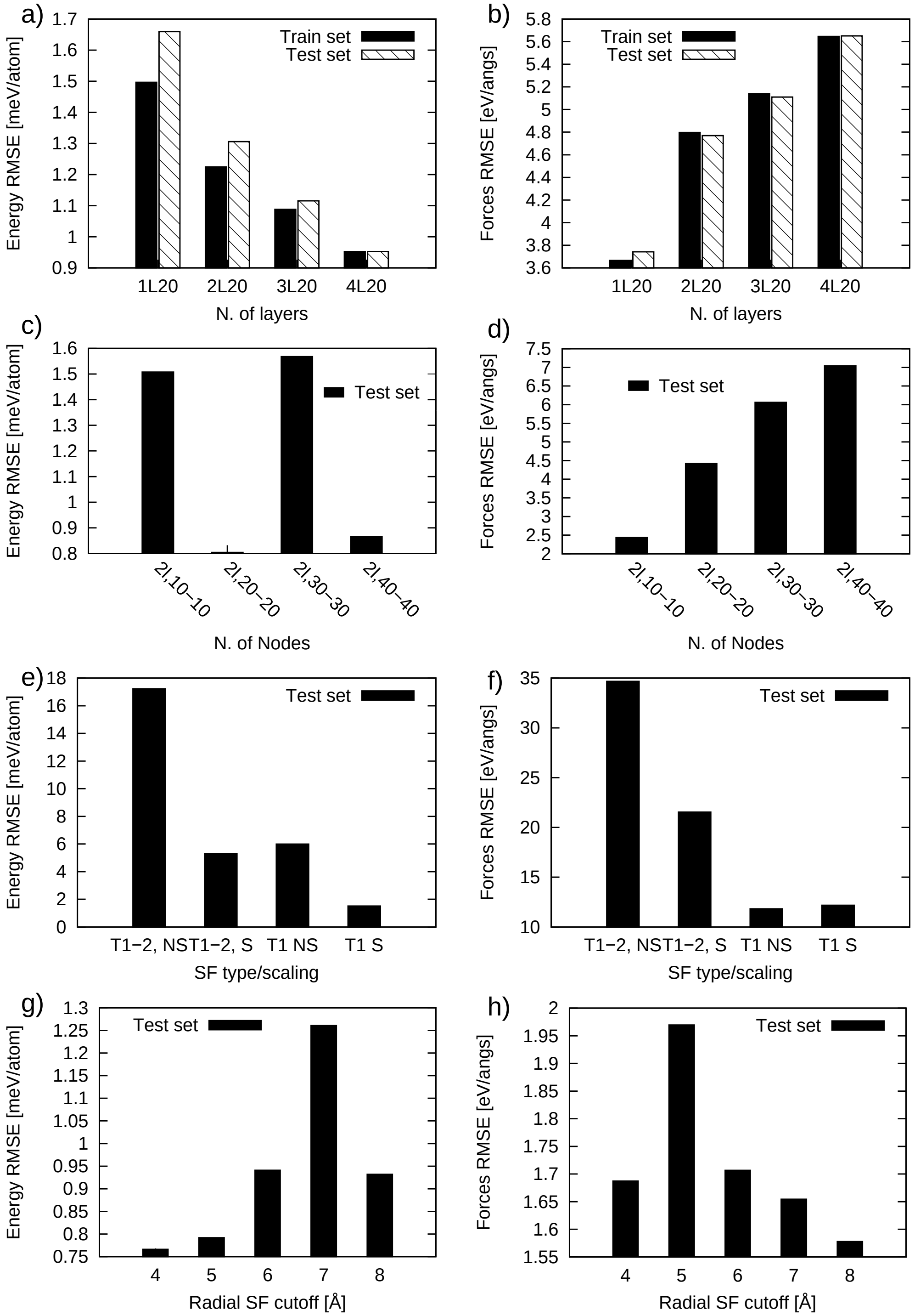

Figure 3.4: Comparison of energy and forces RMSEs for different NN architectures and parameters.

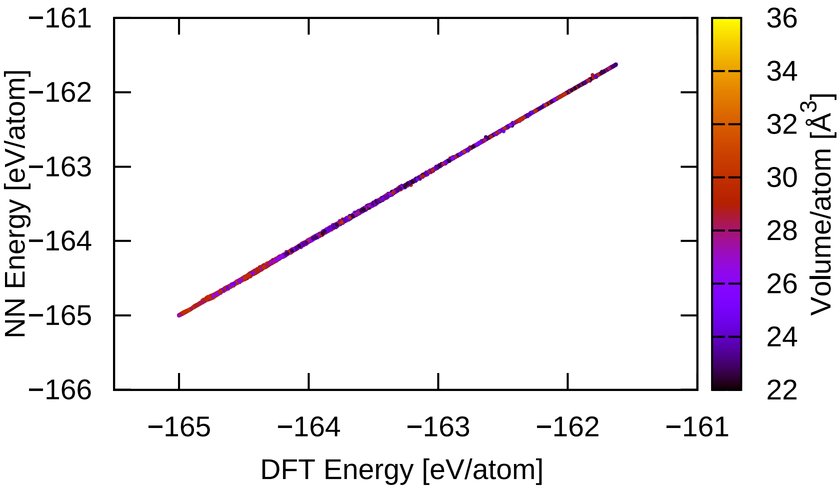

Figure 3.5: Comparison of NN and DFT energies for the whole dataset, including the test set.

3.6 insets). We have found that these configurations correspond to high-energy structures that are not visited in MD simulations carried out in the present work.

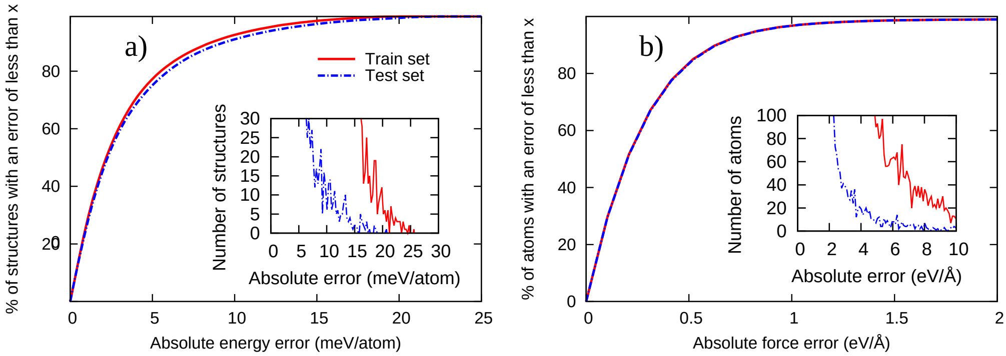

Figure 3.6: Normalized cumulative histograms of the absolute NN errors in training and test data sets for the energies (a) and forces (b). Standard histograms for the same data are presented as insets.

3.2 Structural and Dynamical Properties of Liquid, Amorphous and Crystalline GeTe

In this section we present our results on the validation of our NN potential. We investigated structural and dynamical properties of the bulk phases of GeTe, comparing the NN results with DFT data. Comparison of NN results of the liquid and amorphous phases with experimental data is not discussed here, since the ability of DFT simulations in describing the structural and vibrational properties of GeTe has been demonstrated in a number of previous works [7, 8, 26, 33] and discussed already in section 1.2.

3.2.1 Crystalline phase

The equilibrium geometry of the trigonal phase of crystalline GeTe ($R3m$ space group) was obtained by optimizing all structural parameters consisting of the lattice parameter a, the trigonal angle α and the internal parameter x that assigns the positions of the two atoms in the unit cell, namely Ge at (x,x,x) and Te at (-x,-x,-x) [128]. The residual anisotropy in the stress tensor at the optimized lattice parameter at each volume is below 0.02 kbar. The energy versus volume data were fitted with a Murnaghan equation of state [130]. The theoretical structural parameters of the trigonal phase of GeTe at equilibrium are compared in Tab. 3.1 with experimental data [128] and DFT results obtained with a $12 \times 12 \times 12$ MP k-point mesh in the BZ integration. DFT data are similar to those reported previously [131]. The length of the short and long Ge-Te bonds are also given. The structure of trigonal GeTe can be seen as a distorted rocksalt geometry with an elongation of the cube diagonal along the [111] direction and an off-center displacement of the inner Te atom along the [111] direction, which moves to a distance d from the Ge atom at the vertex as shown in Fig. 3.7a. The energy gained by the off-center displacement is analyzed by varying the distance d at fixed lattice parameters a and α. The resulting energy as a function of d is reported in Fig.3.7 for the NN and the DFT calculations.

We note that the DFT values in Fig.3.7 were not included in the training set but were calculated to investigate the transferability of the NN potential. The double well potential identifies the two possible ferroelectric configurations while the maximum corresponds to an ideal paraelectric configuration.

As a further validation of the potential, we computed the difference in energy between the trigonal phase and an ideal rocksalt phase at their equilibrium volumes at zero temperature that amounts to 44 meV/atom or 55 meV/atom in NN and DFT calculations, respectively.

The elastic properties of trigonal GeTe were investigated by computing the elastic constants from finite deformations of the lattice parameters. The NN and DFT results are compared in Table 3.2. The elastic constants obtained here with the PBE functional are somehow softer than those obtained with the LDA functional in Ref. [132]. The bulk modulus obtained either from the elastic constants or from the equation of state is 34 GPa and 33 GPa for the NN and DFT calculations.

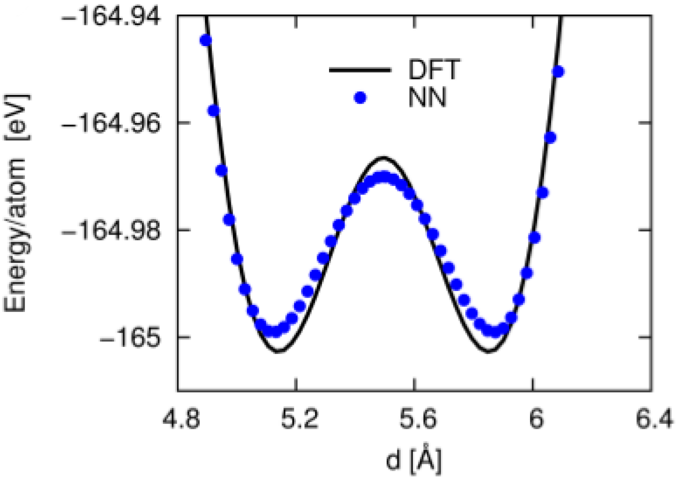

Figure 3.7: Energy of trigonal GeTe as a function of *d* at fixed values of the lattice parameters at the theoretical equilibrium geometry.

Table 3.1: Structural parameters of the trigonal phase of crystalline GeTe from NN and DFT calculations and from the experimental data of Ref.128. The lengths of the short and long bonds are also given.

Structural parameters	NN	DFT	Exp.
a (Å)	4.47	4.33	4.31
α	55.07°	58.14°	57.9°
Volume (Å^3)	55.95	54.98	53.88
x	0.2324	0.2358	0.2366
Short, long bonds (Å)	2.81, 3.31	2.85, 3.21	2.84, 3.17

Table 3.2: Elastic constants (GPa) of trigonal GeTe from DFT and NN calculations.

	c_{11}	c_{12}	c_{13}	c_{14}	c_{33}	c_{44}
DFT	92	18	22	35	40	24
NN	73	10	30	24	36	20

The agreement between NN and DFT results is overall very good, the largest discrepancy being the cell angle and the difference between the short and long Ge-Te bonds which might also be the source of the slight misfit in the elastic constants. The results are overall satisfactory considering that long range Coulomb interactions are not included in our NN potential. In fact, in the case of GeTe, we are faced with the problem of developing a potential suitable to describe both the semiconducting crystalline and amorphous phases as well as the metallic liquid. As a first step toward the development of a NN potential for GeTe, we neglect long-range Coulomb interactions for atoms being separated by a larger distance than the cutoff radius of the symmetry functions (6.88 Å). The resulting "short-ranged" NN potential just consists of atomic energy contributions arising from the local chemical environments, but it is important to note that also short-ranged electrostatic interactions are fully taken into account implicitly. Although long range Coulomb interactions are expected to play a role in the ferroelectric/paraelectric phase change of crystalline GeTe, they are probably less important in the liquid and amorphous phases we are primarily interested in. In fact, the resulting potential is not be able to describe the dielectric response and LO-TO splitting of crystalline GeTe in its ferroelectric phase, but it will be suitable to reproduce structural properties of the liquid, amorphous and crystalline phases and the dynamical properties of the disordered phases. In Fig. 3.8 we show the phonon density of states of crystalline GeTe compared with DFT data [10]. The discrepancy between NN and DFT results is partially due to the lack of long range forces in the NN potential.

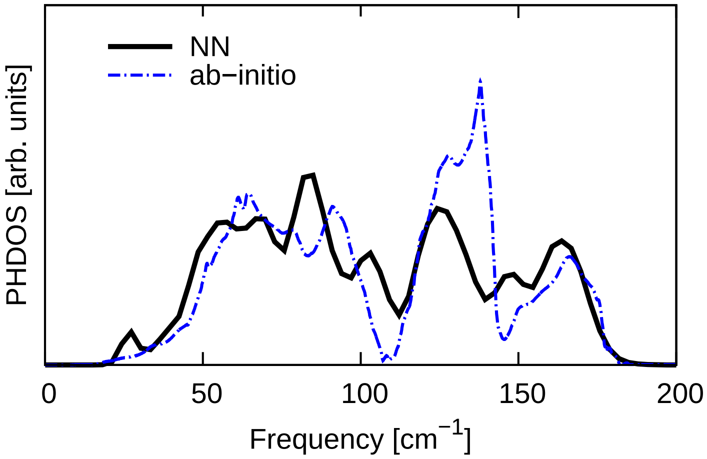

Figure 3.8: Phonon density of states of crystalline trigonal GeTe. DFT data are from Ref. 10.

3.2.2 Liquid phase

The liquid phase of GeTe was simulated by a 4096-atom model at 1150 K. Total and partial pair correlation functions are compared in Fig. 3.9 with results from previous ab initio simulations in a small 216-atom cell at the same temperature [10]. Results from the NN simulations of a 216-atom cell are also reported. The density of 0.0334 atoms/Å^3 is the same for all simulations and corresponds to the value chosen in the ab initio simulations of Ref. 10, which is close to the experimental density of the amorphous phase [129]. Distributions of coordination numbers are reported in Fig. 3.10 as computed by integrating the partial pair correlation functions up to the cutoff shown in Fig. 3.9. Average coordination numbers are given in Table 3.3 while angle distribution functions are shown in Fig. 3.11. The agreement between NN and ab initio data is excellent. The NN results obtained with 216-atom and 4096-atom cells are very close, which demonstrates that structural properties of the liquid can be reliably described by the cells few hundred atoms large used in previous ab initio works [10, 33]. The self-diffusion coefficients computed from NN simulations are also in good agreement with the ab initio results of Ref. 8 as shown in Table 3.4. These latter data refer to simulations at 1000 K to enable a comparison with previous ab initio results obtained at this temperature.

Table 3.3: Average coordination numbers for different pairs of atoms computed from the partial pair correlation functions of liquid GeTe from a NN molecular dynamics simulation at 1150 K with a 4096-atom and a 216-atom cell (cf. Fig.3.9), compared with results from a DFT simulation of a 216-atom cell at the same temperature [10]. The interatomic distance thresholds defined in Fig.3.9 are used.

	With Ge			With Te			Total		
	DFT	NN_{216}	NN_{4096}	DFT	NN_{216}	NN_{4096}	DFT	NN_{216}	NN_{4096}
Ge	1.00	1.11	1.15	2.71	2.78	2.67	3.71	3.89	3.82
Te	2.71	2.78	2.67	0.26	0.28	0.26	2.97	3.07	2.93

Table 3.4: Diffusion coefficient of Ge and Te atoms in the 4096-atom model of liquid GeTe at 1000 K. NN results were obtained from the slope of the mean square displacement versus time. The same values within the figures given here are obtained from the integral of the velocity-velocity autocorrelation function. DFT data of a 216-atom cell at the same temperature are from Ref.8.

	NN	DFT
D_{Ge} (10^{-5} cm^2/s)	4.96	4.65
D_{Te} (10^{-5} cm^2/s)	3.62	3.93

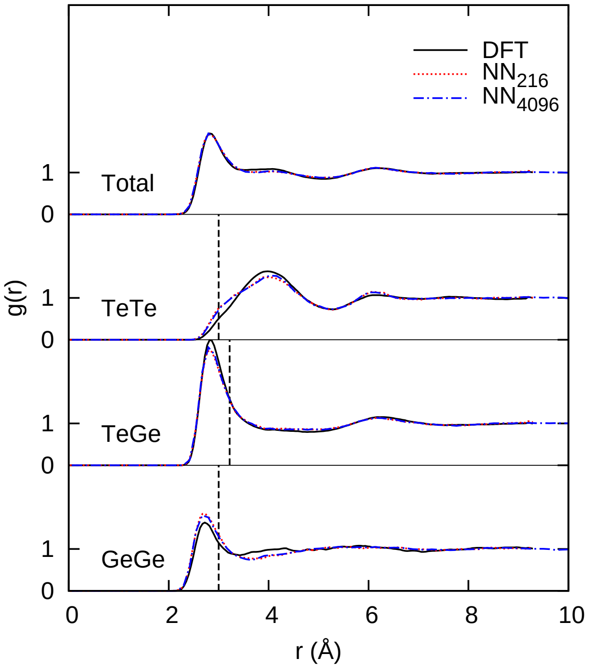

Figure 3.9: Total and partial pair correlation functions of liquid GeTe from a NN molecular dynamics simulation at 1150 K with a 4096-atom and a 216-atom cell, compared with results from a DFT simulation at the same temperature using a 216-atom cell [10]. NN results are obtained by averaging over a NVE run 40 ps long at the average temperature of 1150 K. The vertical lines are the interatomic distance threshold used to define the coordination numbers, 3.0 Å, 3.22 Å, and 3.0 Å for Ge-Ge, Ge-Te and Te-Te bonds, respectively.

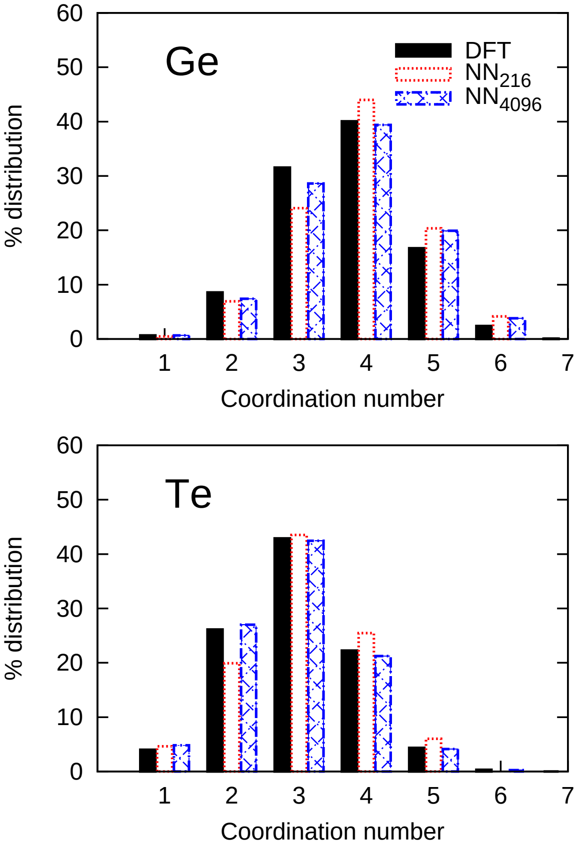

Figure 3.10: Distributions of coordination numbers of Ge and Te atoms in liquid GeTe at 1150 K. Results from NN (4096-atom and 216-atom) and DFT (216-atom) [10] simulations are compared.

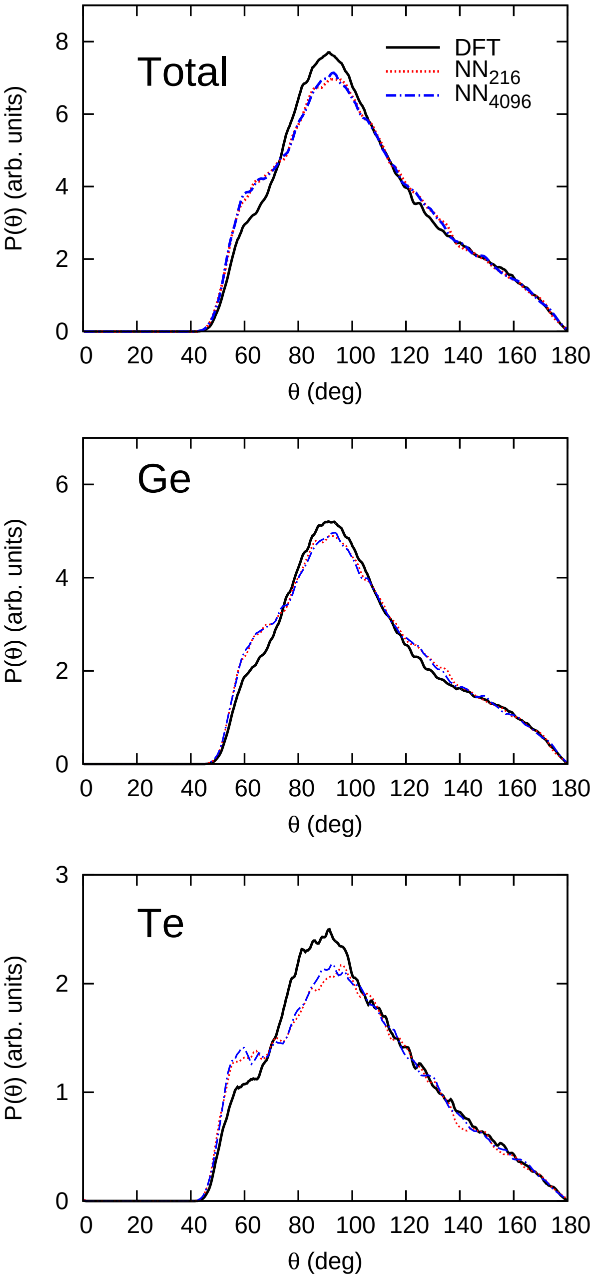

Figure 3.11: Total and partial angle distribution functions of liquid GeTe from a NN molecular dynamics simulation at 1150 K with a 4096-atom and a 216-atom cell, compared with results from a DFT simulation at the same temperature using a 216-atom cell [10]. Partial distributions refer to X-Ge-Y and X-Te-Y triplets (X,Y= Ge or Te).

3.2.3 Amorphous phase

The structural properties of a-GeTe and a-GST have been elucidated recently by ab initio simulations [7, 8, 10, 133]. In these systems Ge and Te atoms are mostly four-coordinated and three-coordinated, respectively. Te atoms are in a defective octahedral-like environment, which resembles the local environment of the corresponding crystalline phases. The majority of Ge atoms are in a defective octahedral environment too, but about one quarter of Ge atoms are in a tetrahedral-like geometry. The presence of homopolar Ge-Ge (and, in the case of GST, Ge-Sb) bonds favors the tetrahedral coordination.

In the following, we compare the structural properties of models of amorphous GeTe generated by NN and DFT simulations. The DFT simulations were performed with the CP2k code. A large NN amorphous model was generated by quenching the 4096-atom liquid model. We also considered ten small 216-atom NN models of a-GeTe to investigate the size of the fluctuations in the structural properties due to the use of a small cell. For sake of comparison, we similarly generated ten 216-atom DFT models of a-GeTe by quenching the melt in 100 ps at the same density of 0.0334 atoms/Å^3. The NN amorphous models were generated by quenching the molten sample from 1150 K to room temperature in 100 ps. Average properties are obtained from a NVE simulation 40 ps long at an average temperature of 300 K. Doubling or even tripling the quenching time (up to 300 ps) does not introduce sizable changes in the structural and vibrational properties of our NN 4096-atom model of a-GeTe. Structural properties are described in Figs. 3.12-3.15. The partial pair correlation functions of our NN models are compared with DFT data in Fig. 3.12. The distribution of coordination numbers and their average values are reported in Fig. 3.13 and Table 3.5 for NN and DFT simulations. Bond angles distribution functions are reported in Fig. 3.14. By decreasing the system size from 4096-atom to 1728-atom one obtains essentially the same results. The NN and DFT data of the small 216-atom cell are averaged over the ten independent models. The properties averaged over the ten NN 216-atom are close to those of the larger 4096-atom NN model. However, in the small cells we observed fluctuations in the structural properties within the ten independent models both for NN and DFT simulations. The fluctuations are slightly larger for the NN models than for the DFT ones.

The agreement between NN and DFT data is overall very good. The largest discrepancy is on the height of the first peak of the Ge-Ge pair correlation function. Another discrepancy with the DFT results is the presence of a small peak at around 60° in the NN angle distribution function due to a very small fraction of three membered rings (see below). The same discrepancy is present in the NN angle distribution function of the liquid (cf. Fig. 3.9). Following previous works [7], we quantified the fraction of Ge atoms in a tetrahedral geometry by computing the local order parameter $q = 1 - \frac{3}{8} \sum_{i>k} (\frac{1}{3} + cos\theta_{ijk})^2$ where the sum runs over the pairs of atoms bonded to a central atom j. $q = 1$ for the ideal tetrahedral geometry, $q = 0$ for the six-coordinated octahedral site, and $q = 5/8$ for a four-coordinated defective octahedral site. The distribution of the local order parameter q for Ge atoms is reported in Fig. 3.15 for different coordination numbers. The q distri-

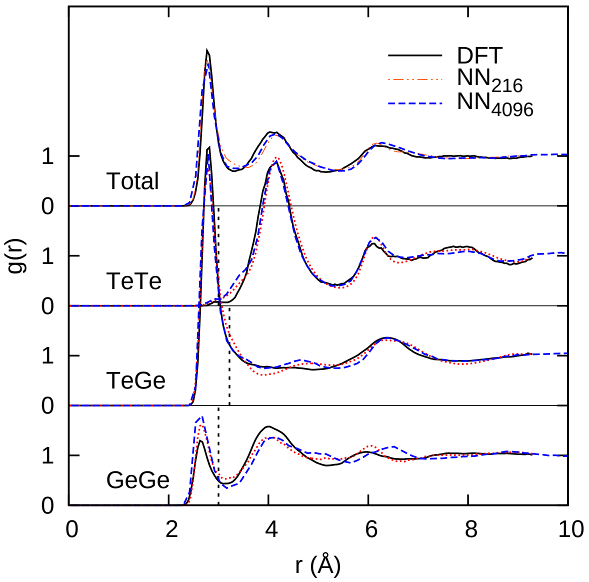

Figure 3.12: Total and partial pair correlation functions of amorphous GeTe from a NN molecular dynamics simulation at 300 K with a 4096-atom and 216-atom cell, compared with results from DFT simulation at the same temperature using 216-atom cells. The vertical lines are the interatomic distance thresholds used to define the coordination numbers, 3.0 Å 3.22 Å and 3.0 Å for Ge-Ge, Ge-Te and Te-Te bonds, respectively. NN and DFT data for the small cell are averaged over ten independent models here and in all the subsequent figures if not specified otherwise.

Table 3.5: Average coordination number for different pairs of atoms computed from the partial pair correlation functions of amorphous GeTe from NN molecular dynamics simulations at 300 K with a 4096-atom and 216-atom cell, compared with results from DFT simulations at the same temperature of 216-atom cells. The interatomic distance thresholds defined in Fig. 3.12 are used.

	With Ge			With Te			Total		
	DFT	NN_{216}	NN_{4096}	DFT	NN_{216}	NN_{4096}	DFT	NN_{216}	NN_{4096}
Ge	0.76	0.78	0.88	3.28	3.31	3.22	4.03	4.09	4.10
Te	3.28	3.31	3.22	0.02	0.04	0.05	3.30	3.35	3.27

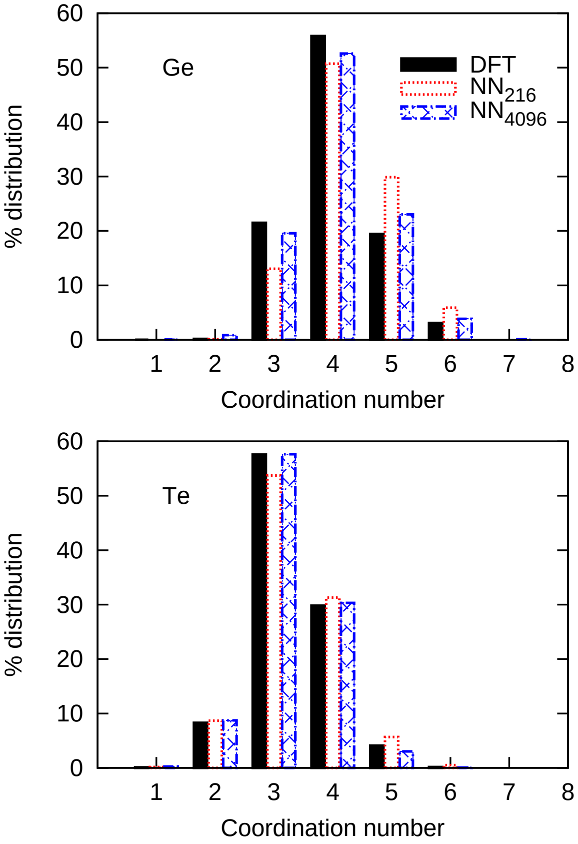

Figure 3.13: Distribution of coordination numbers of Ge and Te atoms in a-GeTe from NN and DFT simulations with 4096- and 216-atom cells. The interatomic distance thresholds defined in Fig. 3.12 are used.

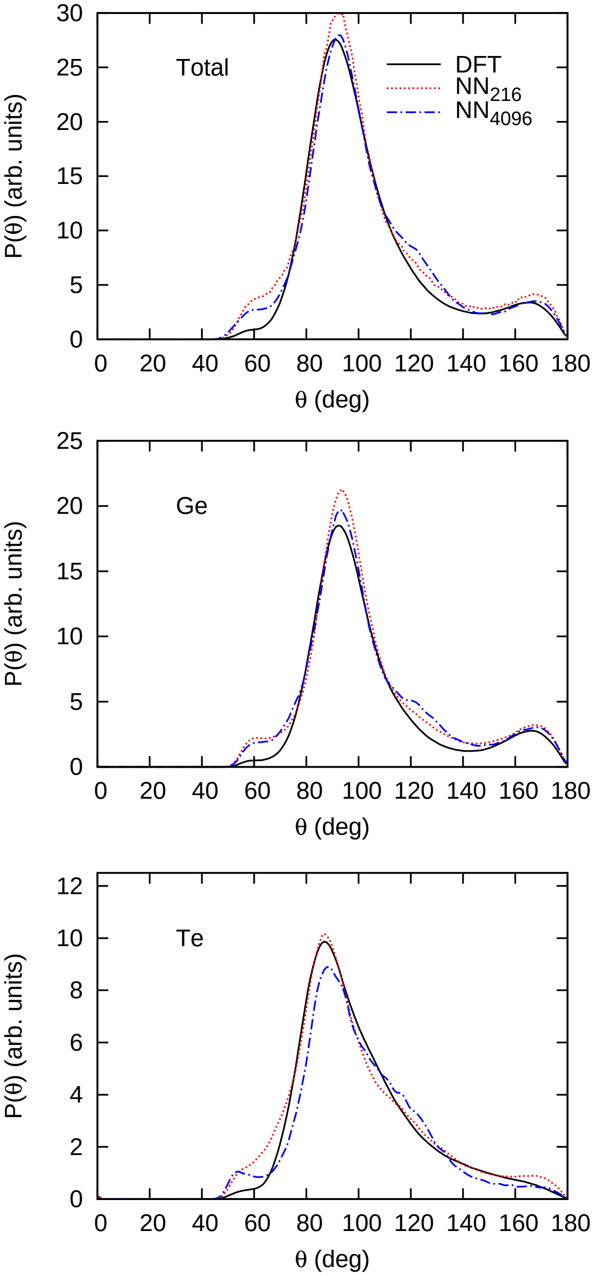

Figure 3.14: Total and partial angle distribution functions of amorphous GeTe from a NN molecular dynamics simulation at 300 K with a 4096-atom and 216-atom cell, compared with results from DFT simulation at the same temperature of 216-atom cells. Partial distributions refer to X-Ge-Y and X-Te-Y triplets (X,Y= Ge or Te).

bution for 4-coordinated Ge is bimodal with peaks corresponding to defective octahedra and tetrahedra. In contrast, the q-distribution for Te does not show any signature of the tetrahedral geometry (cf. Fig. 3.15). We estimated the fraction of tetrahedral Ge atoms by integrating the q-distribution of 4-coordinated Ge from 0.8 to 1. This procedure was demonstrated to provide reliable values for the fraction of tetrahedral Ge from the analysis of the Wannier functions that allow a direct identification of the tetrahedral geometry in terms of the electronic structure [133]. In fact, Ge in tetrahedral sites has four bonding sp^3-like Wannier functions, while Ge in defective octahedra has three p-like bonding Wannier functions and one s-like lone pair. The fraction of tetrahedral Ge atoms for all the ten 216-atom NN models, for the large 4096-atom model and the DFT result average over ten 216-atom models are compared in Fig. 3.16. The average values of the small 216-atom DFT and NN models are both very close to the value for the large 4096-atom NN model (24 %) and also for an intermediate 1728-atom model (22 %).

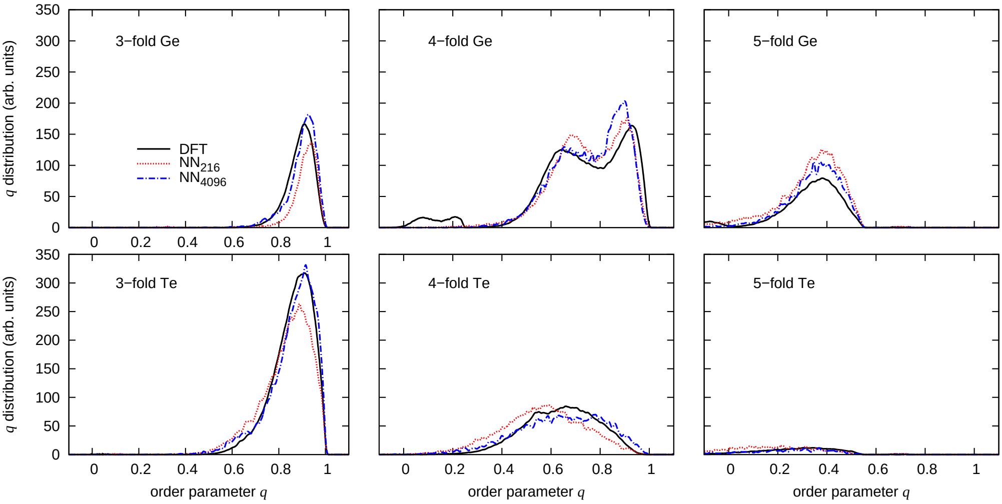

Figure 3.15: Order parameter q for tetrahedricity for 3-, 4-, and 5-fold coordinated Ge and Te atoms in amorphous GeTe from NN molecular dynamics simulations at 300 K with a 4096-atom and 216-atom cells, compared with results from a DFT simulation at the same temperature and 216-atom cells.

Turning now to the medium range order, we report in Fig. 3.17 the distribution of ring lengths computed according to Ref. 32 for the large and small NN models and for the small DFT models. The amorphous phases of GeTe and GST have been shown to display a large concentration of nanocavities [8]. The distribution of the volume of nanocavities computed according to the definition of Ref. 117 and the algorithm of Ref. 118 is compared in Fig. 3.18 for the NN and DFT models. The same scheme for the calculation of nanocavities was applied in previous works on different phase change materials [33]. These comparisons show that the agreement between NN and DFT results is very good for the medium range

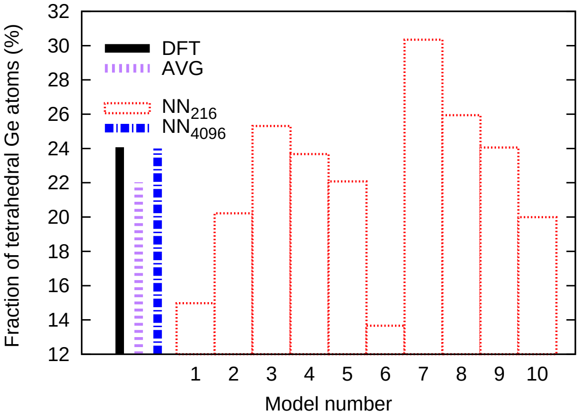

Figure 3.16: Fraction of tetrahedral Ge atoms in amorphous GeTe from ten statistically independent 216-atom models generated with the NN potential, compared with results averaged over ten DFT 216-atom models. The average (AVG) over the ten NN models and from the 4096-atom model are also reported.

order as well.

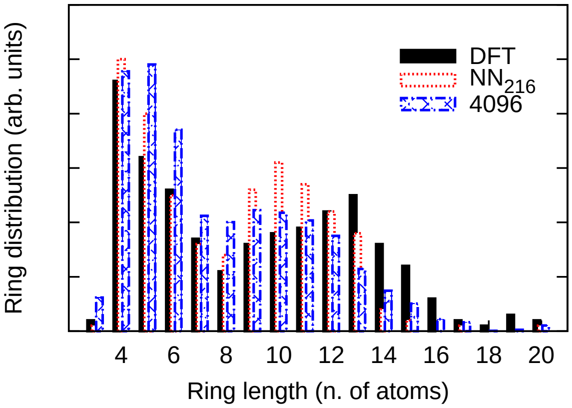

Figure 3.17: Distribution of ring lengths in the 4096- and 216-atom NN models and in the DFT 216-atom models.

We also optimized the density of the amorphous model at zero temperature by interpolating the energy-volume points with a Murnaghan equation of state. We obtained an equilibrium density of 0.03351 atoms Å^{-3} to be compared with the value of 0.03156 atoms Å^{-3} resulting from the DFT equation of state of a 216-atom cell with the BZ integration restricted to the Γ-point [10]. The experimental equilibrium density [129] of a-GeTe is 0.03327 atoms Å^{-3}. The NN and DFT bulk moduli of a-GeTe are 17 GPa and 14 GPa,

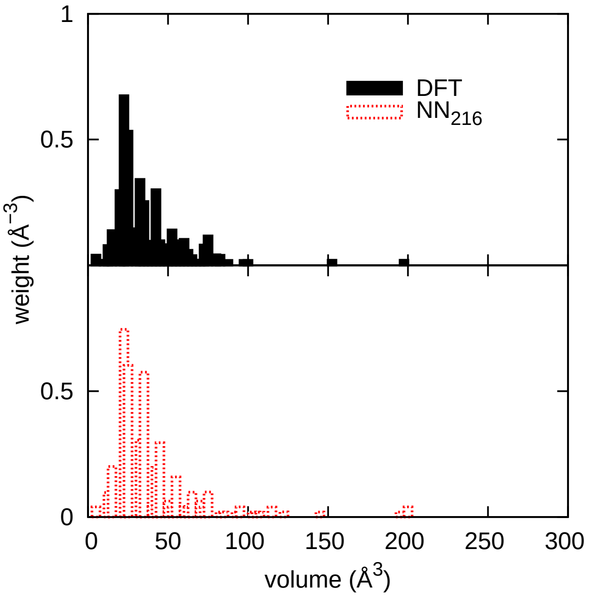

Figure 3.18: Distribution of the volume of cavities in the 216-atom NN and DFT models.

respectively.

Concerning the vibrational properties, the phonon density of states of amorphous GeTe from NN and DFT simulations are compared in Fig. 3.19. Phonon frequencies are computed by diagonalizing the dynamical matrix obtained in turn from the variation of atomic forces due to finite atomic displacements 0.02 Å large. Only phonons with the periodicity of our supercells (Γ-point phonons) were considered. Ab initio phonons are computed in a single 216-atom cell [10] while NN phonons are obtained from either the 4096-atom cell or averaged over ten 216-atom models. Projections on the different type of atoms (Te, Ge in tetrahedral and defective octahedral geometries) are also shown.

In an amorphous material, phonons display localization properties, which depend on frequency. To address this issue and following our previous DFT works [10], we computed the inverse participation ratio (IPR) of the j-th vibrational mode (see section 1.3).

The values of IPR for the NN and DFT models of a-GeTe are reported in Fig. 3.20. The NN potential reproduces the strong localization on tetrahedra of phonons above 200 cm^{-1}. The overall shape and frequency range of the phonon DOS is reasonably reproduced by the NN potential. A discrepancy is present in the relative height of the two main structures at 50 cm^{-1} and 150 cm^{-1}, which, however, might be partially due to the still small size of the 216-atom cell. The size of the fluctuations in the DOS among the 216-atom NN models can be appreciated in Fig. 6.19 in appendix 6.2. The total DOS averaged over the ten models is however close to that of the larger 4096-atom model although differences in the projected DOS are still sizable.

The results discussed in this chapter, on the development of the NN potential and its validation by comparison with DFT results have been published in Ref. [134].

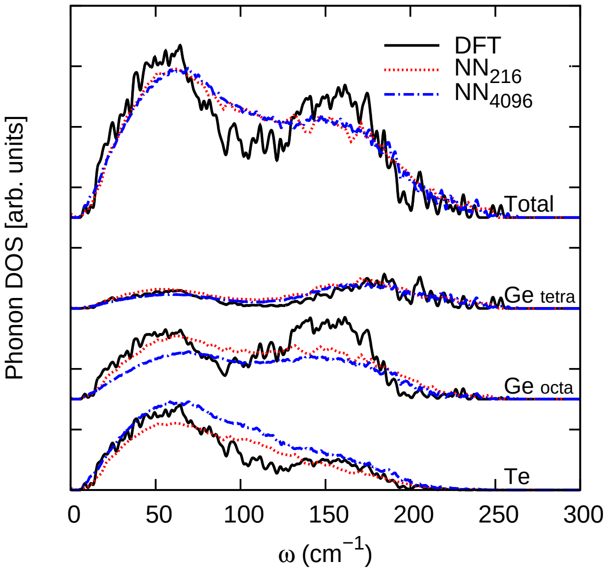

Figure 3.19: Phonon density of states of amorphous GeTe from 4096- and 216-atom NN models. Results for a single 216-atom DFT model [10] are also reported. Only phonons at the supercell Γ point are considered. Projections of the DOS on the different atomic species (Te atoms and Ge atoms in tetrahedral and defective octahedral geometries) are also shown.

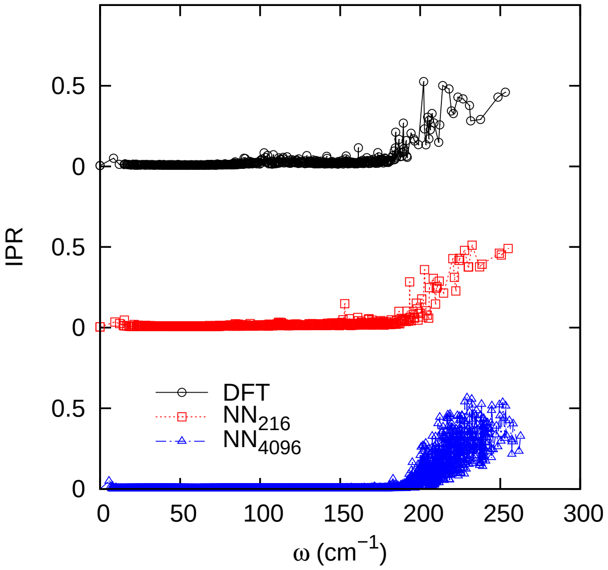

Figure 3.20: Inverse participation ratio of phonons in a-GeTe from the 4096-atom model and a single 216-atom NN and DFT model [10].

Chapter 4

Thermal conductivity of amorphous GeTe

Thermal conductivity tensor $\kappa_{\mu\nu}$ relates the heat current J_μ that flows in a material to the temperature gradient $\frac{\partial T}{\partial x_\nu}$ via Fourier's law as

$$J_\mu = -\sum_\nu \kappa_{\mu\nu} \frac{\partial T}{\partial x_\nu} \tag{4.1}$$

The thermal conductivity of a material is the sum of two different contribution: the electronic thermal conductivity, which depends on the electronic band structure, electron scattering, and electron-phonon interaction, and the lattice conductivity, which depends on phonons and phonon scattering. In this thesis work, we only considered the lattice thermal conductivity, as in amorphous GeTe electrons give a negligible contribution to the thermal conductivity due to the very low electrical conductivity [11]. The lattice thermal conductivity can be computed by means of molecular dynamics with two different approaches, the direct method [135] and the Green-Kubo method [135]. The direct method is a non equilibrium MD method that relies on imposing a temperature gradient across the simulation cell, and it is therefore analogous to the experimental situation. By contrast, the Green-Kubo method approach is an equilibrium MD method that allows to computing the thermal conductivity from energy current fluctuations via the fluctuation-dissipation theorem. We chose to apply the latter approach, which is described in detail in the following. The Green-Kubo method in a general technique to compute response functions in the linear response regime based on the fluctuation dissipation theorem.

The dynamical properties of a system show fluctuations at equilibrium and these fluctuations can be used to determine the desired properties. This is made possible through linear response theory by using relations proven by Green-Kubo (GK) or alternatively - and equivalently- by Einstein. Transport coefficients are defined in terms of the response of the system to perturbations. The fluctuation-dissipation theorem states that the response function is related to the fluctuation of the observable A which is coupled to the external perturbation. Thus, properties are considered as perturbations from equilibrium values. The transport coefficient γ is determined by the autocorrelation function of A via

$$\gamma = \int_0^\infty \langle \dot{A}(t)\dot{A}(0)\rangle dt \tag{4.2}$$

Green-Kubo relations can be derived for many transport coefficients such as diffusivity, shear viscosity, and thermal conductivity. The expression for the latter involves the autocorrelation function of the energy current given by

$$\kappa = \frac{1}{3k_B T^2 \Omega} \sum_{\alpha=1}^{3} \int_0^\infty \langle J_\alpha(0) J_\alpha(t')\rangle dt', \tag{4.3}$$

where α runs over Cartesian coordinates, k_B is the Boltzmann constant, T the temperature and Ω is the volume of the system. The object J in Eq. 4.3 is ensemble dependent. In the microcanonical (NVE) ensemble $J_\alpha(t')$ is the energy current, which can be written as

$$J_\alpha(t') = -\sum_{i=1}^{N} \sum_{\beta=1}^{3} \sigma_{i,\alpha\beta}(t') \cdot v_{i,\beta}(t') = \sum_{i=1}^{N} J_{i\alpha}(t') \tag{4.4}$$

where $v_{i,\beta}(t')$ is the velocity of atom i at time t' and $\sigma_{i,\alpha\beta}(t')$ are the elements of the atomic stress tensor. It is worth to notice that usually there is no unique way of partitioning the total stress into the sum of single-particle contributions. However, in the NN framework the total energy of the system is written as the sum of the energies of individual atoms, so that the definition of an atomic stress follows in a straightforward and consistent manner [124].

Since diffusion processes are negligible in a-GeTe at 300 K, we neglected the convective contribution to the energy flux [136]. To achieve a faster convergence and a better statistical accuracy in the time integral and averages, we can recast Eq. 4.3 by using the Einstein relation [137]

$$\kappa = \frac{1}{6k_B T^2 \Omega} \lim_{t\to\infty} \frac{d}{dt} \sum_{i=1}^{N} \sum_{\alpha=1}^{3} \langle [\epsilon_{i,\alpha}(t) - \epsilon_{i,\alpha}(0)]^2 \rangle, \tag{4.5}$$

with

$$\epsilon_{i,\alpha}(t) - \epsilon_{i,\alpha}(0) = \int_0^t dt' J_{i\alpha}(t'). \tag{4.6}$$

Very long simulation times appear to be needed to sufficiently converge the current-current autocorrelation function [138]. The truncation time in Eq. 4.6 was tested and settled to 40 ps, but long simulations (2 ns) are needed to gain good accuracy in the average $\langle [\epsilon_{i,\alpha}(t) - \epsilon_{i,\alpha}(0)]^2 \rangle$ over the initial state ($t = 0$) in Eq. 4.5. We used four independent

4096-atom models of a-GeTe generated by quenching from the melt to 300 K in 100 ps. The density was fixed to the value of 0.0334 atoms/Å^3 very close to the experimental one for the amorphous phase [129]. The models were equilibrated at 300 K by stochastic velocity rescaling [139].

The value of κ resulting from the GK formula (Eq. 4.2) at 300 K is 0.27±0.05 W/(mK), where the error bar is the sum of two contributions, one from the uncertainties in the linear fit of Eq.4.5 (± 0.048 W/(mK), see Fig.4.1) and a second due to the spread in the value of κ obtained from the four different models (± 0.0087 W/(mK)).

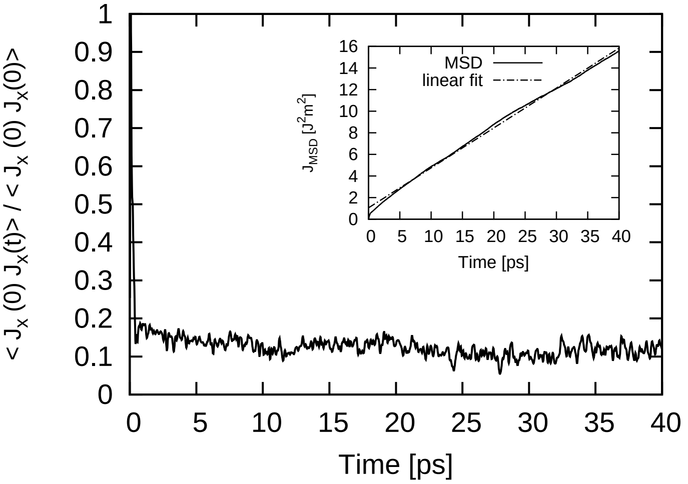

Figure 4.1: Normalized autocorrelation function of the energy current operator. Inset shows the MSD of J according to Eq. 4.5.

Our value of κ is close to those measured experimentally for $\mathrm{Ge}_x\mathrm{Te}_{1-x}$ and GeSbTe alloys [11], all in the range 0.1-0.3 W/mK, including $\mathrm{Ge}_{0.20}\mathrm{Te}_{0.80}$ (0.19 W/mK) [12]. An experimental value of κ is available also for a-GeTe [140] which, however, is about 2.5 W/(mK), one order of magnitude larger than our result and than other experimental data on very similar systems. The origin of this discrepancy is unclear and suggests the need of new experimental measurements on a-GeTe.

The value of κ obtained by MD simulations (either via the Green-Kubo approach or the NEMD direct method) accounts for the whole value of the lattice thermal conductivity, that is, it collects all contributions coming from the vibrational modes of the systems. However, as discussed in section 1.3 in the case of amorphous materials, there exist different kind of phonons, each one contributing to the total value of κ by a different mechanism. It is therefore important to investigate further the microscopic origin of κ to identify which and to what extent different phonons contribute to thermal conductivity.

As discussed in section 1.3 propagating and non-propagating (diffusons) phonons contribute to the thermal conductivity of an amorphous material. In the following we discuss separately the contributions from the two types of phonons.

4.1 Propagating phonons

The Peierls-Boltzmann theory allows one to compute the contribution to the thermal conductivity of propagating phonons by applying the Boltzmann Transport Equation (BTE), under certain assumptions, to a phonon gas. We begin by rewriting Eq. 4.1 as

$$\vec{J} = -\vec{\kappa} \cdot \nabla T \tag{4.7}$$

The BTE can be used to track the time evolution of the positions and velocities of a system of particles. In the case of a gas of phonon in steady state to first order in the temperature gradient, the BTE for a specific phonon mode takes the form [141]

$$\vec{v}_g \cdot \nabla T \frac{\partial f_t}{\partial T} = \left(\frac{\partial f_t}{\partial t} \right)_{coll} \tag{4.8}$$

where f_t and $\vec{v}_g$ are the phonon distribution function and group velocity vector and $\frac{\partial f_t}{\partial t}_{coll}$ is the collision term. The left hand side of Eq. 4.8 describes a system of non interacting phonons. On the right-hand side, the collision term accounts for all possible mechanism of phonon interaction and is generally quite complex. The main challenge in working with BTE is in specifying and evaluating an expression for the collision term. The single relaxation time approximation (SRTA) is commonly used to make the BTE tractable. In SRTA, the phonon distribution functions and the collision term are written as

$$\begin{aligned} f_t &= f + f_t' \\ \left(\frac{\partial f_t}{\partial t} \right)_{coll} &= -\frac{f'}{\tau} \end{aligned} \tag{4.9}$$

where f is the Bose-Einstein equilibrium distribution function $f_i = (e^{\beta \hbar \omega_i} - 1)^{-1}$, f' is the change of the distribution function from equilibrium, and τ is the phonon relaxation time (lifetime). Now, using the SRTA we write the following expression for distribution about equilibrium starting from Eq. 4.8

$$\vec{v}_g \cdot \nabla T \frac{\partial f_t}{\partial T} = -\frac{f'}{\tau}$$

which at the first order in ∇T leads to

$$f' = -\vec{v}_g \cdot \nabla T \frac{\partial f}{\partial T} \tau \tag{4.10}$$

In the harmonic approximation, the volumetric phonon specific heat c_i is

$$c_i = \frac{\hbar\omega_i}{V}\frac{\partial f}{\partial T} \tag{4.11}$$

The net flux due to phonon motion is given in terms of f' by [142]

$$\vec{J} = \frac{1}{V}\sum_i \hbar\omega_i \vec{v}_g f' \tag{4.12}$$

where V is the system volume and the sum runs over the phonon branches. By pluggin Eq. 4.10 into Eq. 4.12 we obtain the following expression for the thermal conductivity

$$\begin{aligned}
\vec{\kappa} &= -\frac{\vec{J}}{\nabla T} \\
\vec{\kappa} &= -\frac{1}{\nabla T}\frac{1}{V}\sum_i \hbar\omega_i \vec{v}_g f' \\
\vec{\kappa} &= -\frac{1}{\nabla T}\frac{1}{V}\sum_i \hbar\omega_i \vec{v}_g \cdot \left(-\vec{v}_g \cdot \nabla T \frac{\partial f}{\partial T}\tau\right) \\
\vec{\kappa} &= \sum_i c_i \vec{v}_g^2 \tau
\end{aligned} \tag{4.13}$$

We define the mean free path Λ of a certain phonon as $\Lambda = v_g \cdot \tau$. The Peierls-Boltzmann theory is very successful for crystalline matter. For perturbations varying slowly [143] in space and time, the only limitations of the Boltzmann approach is either (a) that each i-th phonon should have a sufficiently long mean free path so that its wave vector $\vec{q}$ can be defined, i.e. $\Lambda_i >> \lambda_i = 2\pi/|\vec{q}|$, or alternately (b) if the wave vector $\vec{q}$ is impressed externally, then the resulting oscillation should have a long enough lifetime (or dephasing time) τ_i so that its frequency ω_i is well defined, i.e. $\tau_i >> T_i = 2\pi/\omega_i$. Amorphous materials have usually only a small bunch of phonons which satisfy both criterions, typically very long wavelength acoustic-like modes. Thus, the Peierls-Boltzmann model can only deal with these particular phonons, that according to the classification we presented in section 1.3 would be propagons. Thus, we need three ingredients in order to compute κ_{BTE}: the specific heat per unit volume c_i, the groups velocity v and the lifetime τ. Specific heat can be easily computed by Eq.4.11. In order to obtain group velocities, we need to define them in the case of an amorphous system. In lattice dynamics calculations, the dynamical matrix $\mathcal{D}$ has been obtained by computing derivatives of forces acting on atoms as finite differences. $\mathcal{D}$ was computed and diagonalized to obtain egienvectors and eigenvalues at the Γ point ($\vec{q} = 0$) of the supercell Brillouin zone and for several small, finite $\vec{q}$ points. In fact, in a supercell the group velocities at the Γ point are zero, except for the three acoustic modes corresponding to $\omega = 0$. However, it is often possible that a number of low frequency modes retain propagating character with $\vec{v}_g \neq 0$ for $\vec{q}$ very close to Γ. Those modes are propagons, for which an effective group velocity can be defined without defining the limit for $\vec{q} \to 0$. We differentiated the dispersion curves over a uniform grid of 18 points within a radius of 0.08 Å^{-1} around the Γ point. Phonon lifetimes are obtained from the autocorrelation functions of the eigenmodes as $\tau_j = 2 \cdot \int_0^\infty \frac{\langle E_j(t) E_j(0)\rangle}{\langle E_j(0) E_j(0)\rangle} dt$ where the kinetic energy of mode j is obtained from the eigenmode amplitude as $E_j(t) = \frac{1}{N} |\sum_{i=1}^N \sqrt{M_i} \ \ \mathbf{e}(j,i) \cdot \vec{v}_i(t)|^2$. The truncation time in the integral above was settled to 20 ps, but a long 2 ns MD trajectory was needed to achieve convergence in the average over the initial state ($t = 0$).

Fig. 4.2 demonstrates that the single relaxation time approximation can be safely applied at least for the majority of a-GeTe phonons. In fact, the energy of the normal mode decades with an oscillatory behavior following an exponential trend that clearly indicates that phonons have a single relaxation time. The resulting group velocities, lifetimes and mean free paths ($\lambda_j = v_{gj} \cdot \tau_j$), calculated as described above, are reported in Fig. 4.3. One can notice that in a-GeTe essentially all the modes have mean free path in the range 1-10 Å i.e. shorter than the average interatomic distance. In fact, group velocities (see panel a) of Fig. 4.3) are really small, of the order of 10^2 m/s, along the whole vibrational spectrum. Only a bunch of acoustic-like phonons display group velocities of the order of 10^3 m/s. In amorphous silicon [15], the fraction of phonons with group velocities of the order of 10^3 m/s, i.e. of the order of the speed of sound, is actually noticeable, as can be expected since silicon is much stiffer than GeTe, and phase materials in general, even in the amorphous phase. Above 200 cm^{-1}, group velocities become even smaller, since phonons beyond that frequency are strongly localized and simply do not propagate at all. Lifetimes are also one order of magnitude smaller with respect to what is obtained in the case of amorphous silicon [15]. In summary, both of the group velocities and the lifetimes are really small, leading to a very small value of κ_{BTE}=0.01 W/mK, which is one order of magnitude smaller than the total value of κ obtained from EMD. Therefore propagating phonons do not contribute sizably to the thermal conductivity of a-GeTe.

4.2 Diffusons

In section 1.3, we introduced diffusons. They are extended, delocalized and not propagating modes. Thus, their contribution to thermal conductivity cannot be computed by BTE or anharmonic hopping models. In the early nineties Allen and Feldman proposed a theory that specifically takes into account diffusons. The theory applies when disorder is sufficient to make the majority of states to propagate very little, but insufficient to make all the states localized. Also it is necessary that the material is stiff enough or the temperature low enough that the harmonic approximation is applicable. In this model, localized states contribute no heat current (anharmonic terms are needed for hopping). Nevertheless, significant heat currents are carried by diffusons. The basic idea is that the heat current operator has off-diagonal matrix elements J_{ij} between the harmonic for foamy systems. The starting point is the Kubo formula of Eq. 4.3. For a disordered harmonic solid the exact many-body states are simply the various ways of occupying the 3N harmonic-oscillator states. If one writes down the heat current operator in terms of these

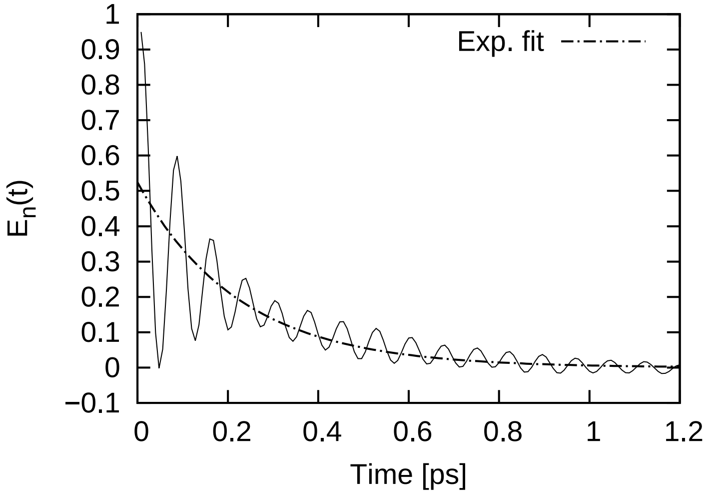

Figure 4.2: Decay of the energy of a normal mode at ~150 cm^{-1} as a function of time.

oscillator states [40], it is possible to recast the Kubo formula Eq. 4.3 as

$$\kappa_{\mu\nu}(\omega) = \frac{\pi\Omega}{T}\sum_{i,j}\frac{f_i - f_j}{\hbar(\omega_i - \omega_j)}(J_\mu)_{ij}(J_\nu)_{ji}\delta(\omega_i - \omega_j - \omega) \tag{4.14}$$

where ω represents the frequency of an applied temperature gradient, J_{ij} are the off diagonal matrix elements of the heat current operator defined in Ref. 40, and f_i is the equilibrium Bose-Einstein occupation of the i-th exact oscillator mode. Eq. 4.14 is the phonon analog of the Kubo-Greenwood approach for electrical conductivity of disordered metals. The factor $(f_i - f_j)/(\omega_i - \omega_j)$ becomes $-\frac{\partial f}{\partial \omega_i}$, and this relates to the specific heat described in Eq. 4.11 which leads to

$$\kappa = \sum_i c_i \frac{1}{3}\sum_{\alpha=1}^{3} D_{\alpha i} \tag{4.15}$$

where D_i is the mode diffusivity, temperature independent, defined in a microscopic way for an harmonic disordered solid as

$$D_{\alpha j} = \frac{\Omega^2}{8\pi^2\hbar^2\nu_i^2}\sum_{n\neq j}|\langle \mathbf{e}_j|J_\alpha|\mathbf{e}_n\rangle|^2\delta(\nu_j - \nu_n) \tag{4.16}$$

Here $\langle \mathbf{e}_n|J_\alpha|\mathbf{e}_j\rangle$ are the matrix elements of the α Cartesian component of the energy flux operator between two harmonic normal mode $\mathbf{e}_n$ and $\mathbf{e}_j$ with frequencies ν_n and ν_j [40].

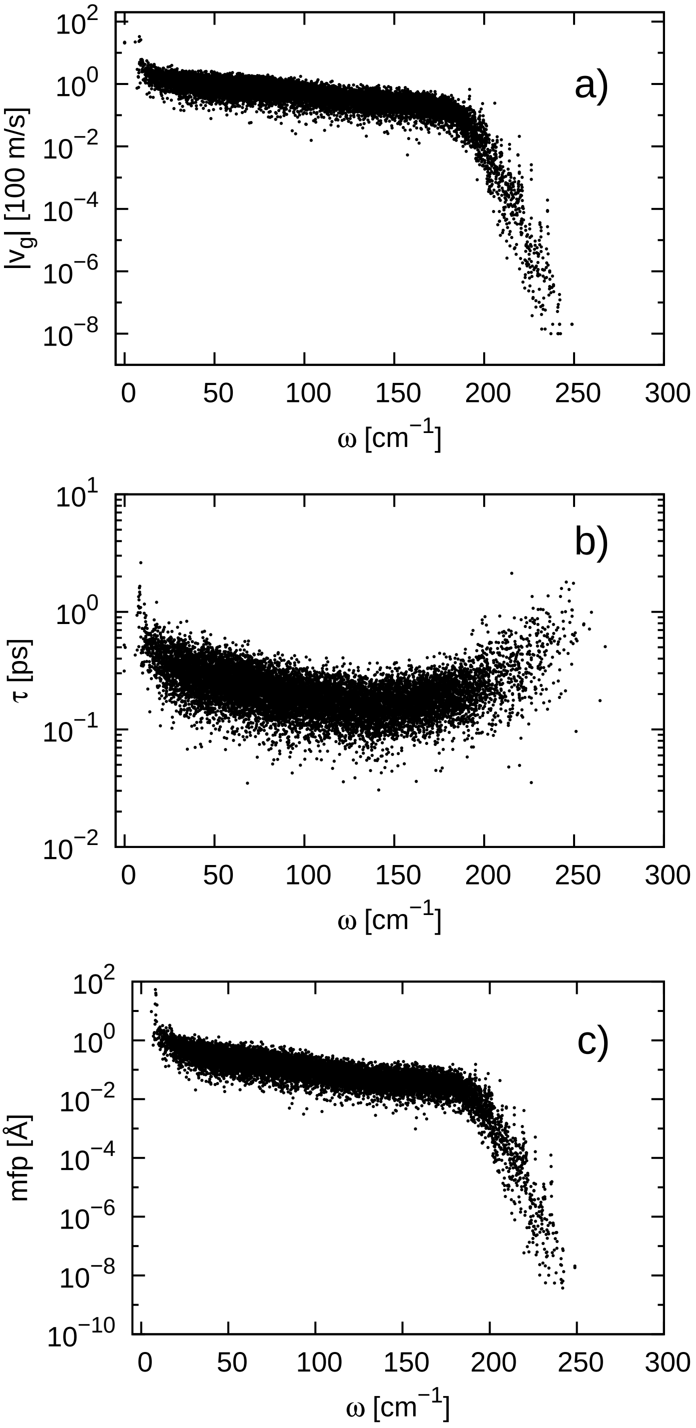

Figure 4.3: a) Group velocities b) lifetimes and c) mean free paths of normal modes in a-GeTe.

In periodic systems, the delta function if Eq. 4.16 must be replaced by a Lorentzian function of a certain width η greater than the spacing between the vibrational levels. The choice of η has a sizable effect on the value of κ obtained in this way, and must be chosen with care. In principle, $\eta \to 0$ when the volume $\Omega \to \infty$, but in practice one has to choose a value of η that allows the overlap of vibrational modes within a tiny frequency window. In any case the choice of η rises an error that must be taken into account. In the case of a-GeTe we chose a broadening δ between 0.1 and 1 cm^{-1}. In Fig. 4.4 we report the diffusivity of each normal mode of a-GeTe for the two choices of δ. One can notice how the width affects the degree of overlap between the normal modes, thus increasing or decreasing their diffusivity. In any case, diffusivity in a-GeTe rapidly decreases as the frequency rises.

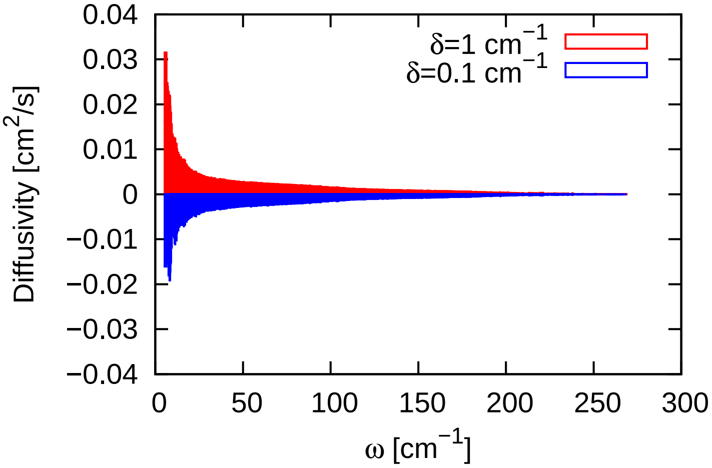

Figure 4.4: Diffusivity of each normal mode of a-GeTe with two different choices of Lorentzian broadening δ.

The cumulative contribution of propagating (κ_{BTE}) and non-propagating modes (κ_{AF}) to the thermal conductivity as a function of the spectral range of the vibrational modes are reported in Fig. 4.6. The error bar, depicted by the dashed area in Fig. 4.6, rises from the uncertainties in our choice of the spreading δ in the range 0.1-1 cm^{-1}. The κ_{BTE} term is negligible compared to κ_{AF} and their sum is very close to the EMD result which indicates that the decomposition $\kappa \simeq \kappa_{BTE} + \kappa_{AF}$ captures essentially all the contributions to the thermal conductivity at 300 K. The specific heat in the expression for κ_{AF} has already reached the classical value at 300 K. Therefore κ_{AF} will not change by increasing temperature above 300 K (see Fig. 4.5).

The cumulative value of κ_{AF} is converged at about 200 cm^{-1} because phonons at higher frequency are mostly localized on tetrahedra and do not contribute to κ_{AF}. On the other hand, the fact that κ_{AF} essentially accounts for the whole value of κ obtained from

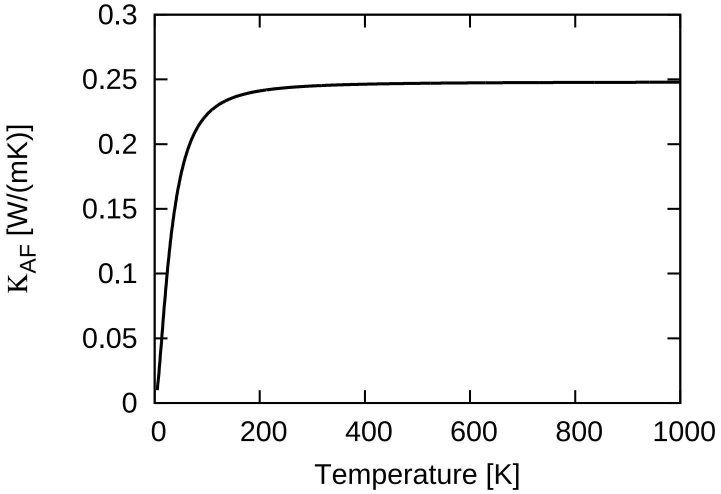

Figure 4.5: Contribution to the thermal conductivity of a-GeTe due to diffusons as a function of frequency. All phonons are populated already at room temperature. $\delta = 0.1$ cm^{-1}.

EMD means that energy transfer among localized modes due to anharmonic interactions are negligible, at least at 300 K. The lack of propagating modes with long mean free path is also responsible for a relatively fast convergence of κ with system size. Indeed the value of κ_{AF} of a smaller 1728-atom model is 0.24±0.01 W/(mK), i.e. almost the same value obtained for the larger 4096-atom model. This result suggests that confinement effects of phonons in nanoscaled GeTe devices are negligible. The lack of propagating modes with long mean free path can be traced back to the rich topology of the amorphous network in a-GeTe which displays a coexistence of tetrahedra, defective octahedra and the presence large nanovoids [8] suitable to scatter phonons and to enhance the fraction of non-propagating, quasi-stationary modes with respect to propagating modes with large group velocities.

An approximate expression for the contribution of diffusons to thermal conductivity under the assumption of a Debye-like phonon DOS was proposed by Cahill *et al.* [144]. In the high temperature limit, where all the modes are thermally excited, which is the case for a-GeTe at 300 K, the predicted minimum thermal conductivity according to Ref. 144 model is $\kappa_{min} = \frac{3}{2} k_B (\frac{\pi}{6})^{\frac{1}{3}} \rho^{\frac{2}{3}} v_s$, where ρ is the atomic density and v_s is the speed of sound taken as an average over longitudinal and transverse modes: $3v_s$=v_L + 2 v_T. In our case, v_s = 2.1 km/s which gives κ_{min} =0.36 W/mK, close the value we obtained for κ_{AF}.

In conclusion, we used our NN interatomic potential to investigate thermal transport in the amorphous phase of the GeTe compound. We computed the thermal conductivity κ from equilibrium molecular dynamics and the Green-Kubo formula. The resulting value κ= 0.27±0.05 W/(mK) is within the range of values measured for several other

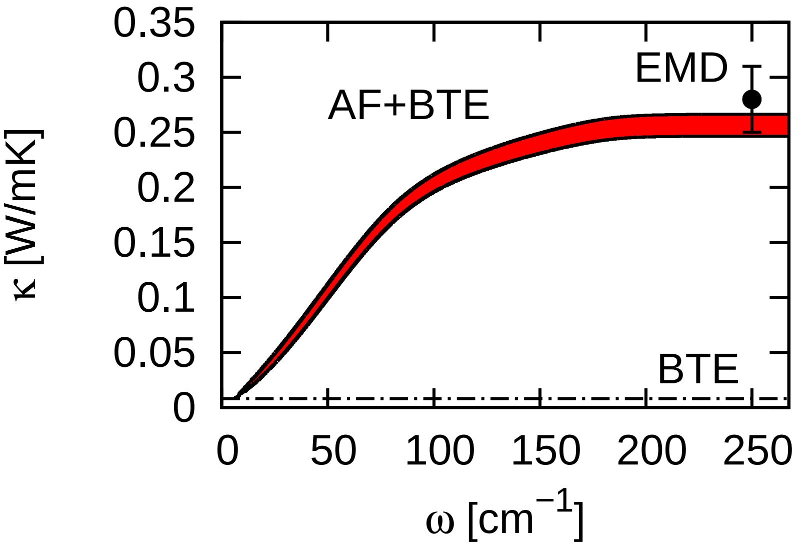

Figure 4.6: Cumulative thermal conductivity of a-GeTe as a function of frequency. The value of κ at ω is due to all phonons with frequency below ω.Contribution from propagating modes (κ_{BTE}) is shown together with his sum (κ_{AF+BTE}) with the contribution from non-propagating modes (κ_{AF}) and the result obtained from equilibrium MD simulations (EMD). The width of the AF+BTE function corresponds to the error bar.

GeSbTe phase change materials. The calculation of phonons group velocity and phonon lifetimes reveal that essentially all phonons have short mean free path in the range 1-10 Å. The contribution to the thermal conductivity by propagating modes is thus negligible. The thermal conductivity is in fact mostly due to non-propagating, delocalized modes (diffusons) which can transport heat according to the theory developed by Allen and Feldman [40]. Our results thus suggest that the bulk value of κ measured for instance in films can indeed be used to model the thermal transport of GeTe also in nanoscale devices, as the contribution from propagating modes that may endure ballistic transport at the scale of 10-20 nm is negligible. Phase change ternary alloys GeSbTe have all very similar lattice thermal conductivity with values [11] close to what we have found for GeTe. The phonon spectrum and the structure of amorphous GeTe are also very similar to those of other GeSbTe compounds with coexistence of tetrahedra and defective octahedra and the presence of nanovoids [8, 10, 33]. It is therefore conceivable that the absence of phonons with long mean free path and the dominance of non-propagating, delocalized modes in the thermal transport found here for GeTe might be extended to other materials in the same class. The results discussed in this chapter, concerning the investigation of thermal transport in the amorphous phase of GeTe, have been published in Ref. [145].

Chapter 5

Viscosity and atomic mobility in the Supercooled Liquid and Overheated Amorphous Phases

As discussed in section. 1.4, the properties of the supercooled liquid control the speed of crystallization. It has been proposed that the higher crystallization speed of phase change materials is due to a high fragility of the supercooled liquid phase. However no experimental data on the viscosity of the supercooled liquid are available. In this work we have addressed this problem by means of NN molecular dynamics which allowed us to compute independently viscosity and diffusivity in large simulation cells. A breakdown of the Stokes-Einstein relation close to the glass transition has been found. We also studied hysteretic effects in the glass transition by analyzing the properties of the overheated amorphous phase.

5.1 Melting temperature

To study the properties of the supercooled liquid, we first assessed the ability of the NN potential, and thus of the underlying DFT-PBE framework, to reproduce T_m. The melting temperature was computed according to the framework described in section 2.3. Thermodynamic integration yielded T_m=1001 K, a number very close to the experimental value at normal pressure of 998 K [21]. To obtain T_m, we first computed the difference in the Helmholtz free energy F between the NN system and a reference system for which an analytic expression for F is known, at a given temperature T' and density ρ'. Namely

$$F_{NN}(T',\rho') - F_{ref}(T',\rho') = \int_0^1 d\lambda \langle U(\lambda)\rangle, \tag{5.1}$$

where the average is taken over a MD simulation with the potential $U(\lambda) = \lambda U_{NN} - (\lambda - 1)U_{ref}$. The temperature and density were set to the experimental values at the melting point at normal conditions [21]. As discussed in section 2.3.1, the reference system was chosen as an Einstein crystal for the solid and a Lennard-Jones fluid for the liquid. In the next step, the chemical potentials were evaluated by integrating the free energy as a function of density (see Eq. 2.64). By equating the chemical potential of the two phases one obtains a transition pressure of -0.44 GPa at the chosen temperature T′=998 K. From the calculated Clausius-Clapeyron equation (dT/dP=6.85 K/GPa from the calculated ΔS=$\Delta E/T$ and ΔV on the theoretical melting line at T=998 K) we then obtained the theoretical melting temperature at normal pressure which is T$_m$=1001 K. $\langle U \rangle$ as a function of λ for the two phases is shown in Fig. 5.1 and the chemical potential of the two phases as a function of pressure at T'=998 K is shown in Fig. 5.2.

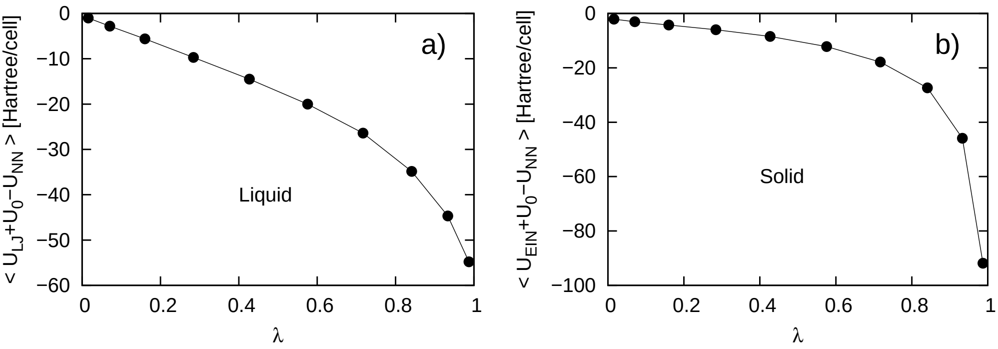

Figure 5.1: Ensamble average of the total potential energy of the system as a function of the mixing parameter λ (see section 2.3 for details). a) Liquid phase of GeTe with a Lennard Jones liquid as reference system b) Crystalline phase of GeTe with an Einstein crystal as reference system.

5.2 Supercooled liquid

We analyzed the properties of the supercooled liquid below T$_m$ by computing independently the viscosity η and the self diffusion coefficient D in microcanonical MD simulations. The volume of the supercooled liquid was scaled with temperature according to the calculated thermal expansion coefficient α =4.73$\cdot$ 10^{-5} K^{-1}. At T_m α is little dependent on temperature [21]. We scaled the temperature from 1000 K to 500 K in eleven steps. At each step the system is equilibrated for 25 ps with the thermostat. Overall the system is thus quenched from 1000 K to 500 K in 250 ps. At each temperature statistical averages are then collected on longer microcanonical simulations up to 2 ns long for the calculation of the viscosity as discussed below.

We first computed D from the atomic mean square displacement (MSD) as

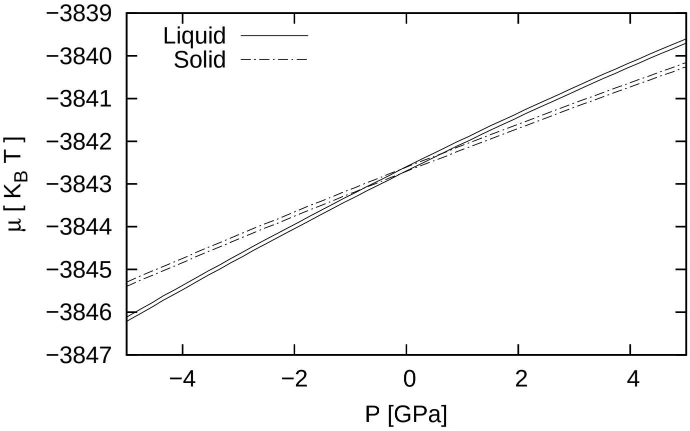

Figure 5.2: Chemical potentials of liquid and solid GeTe obtained by integrating Helmoltz free energies of each phase. Intersection of the two curves is the coexistence point in the pressure/temperature space that assigns the melting temperature at a given pressure.

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{\partial \langle |\vec{r}_i(t) - \vec{r}_i(0)|^2 \rangle}{\partial t} \tag{5.2}$$

where $\vec{r}_i(t)$ is the position of atom i at time t. The average $\langle \ldots \rangle$ is over atoms and over the initial times in $\vec{r}_i(0)$ to improve statistical accuracy. D can be computed on a timescale of 50 ps on which the system does not crystallize at the temperatures we considered. The MSD as a function of time for selected temperatures are shown in Fig. 5.3.

The values of D as a function of temperature are reported in Fig. 5.4a. D is still of the order of 10^{-6} cm^2/s at the lowest temperature of 505 K considered here; it follows an Arrhenius behavior from T$_m$ to 505 K. The activation energy is 0.220 ± 0.002 eV, a value much lower than the activation energy of 1.76 eV obtained from the Arrhenius dependence of viscosity in GST measured in the temperature range 333-373 K (probably below T$_g$) [146]. The ratio between the self-diffusion coefficient of Ge and Te (D_{Ge}/D_{Te}) increases by decreasing temperature as shown in Fig. 5.5. The values of D obtained from the NN simulations were also validated by direct DFT molecular dynamics simulations at few selected temperatures with a small 216-atom cell (cf. Fig. 5.4a). The DFT result at 1000 K is equal to that previously obtained in Ref.[8] with the same cell and functional used here. We checked that the change of volume with temperature has a little effect on the diffusion coefficient as shown in Fig. 5.4a which also reports the values of D as a function of temperature once the density is fixed to the value at the melting point.

We then computed η between T$_m$ and a temperature T*=700 K which turned out to

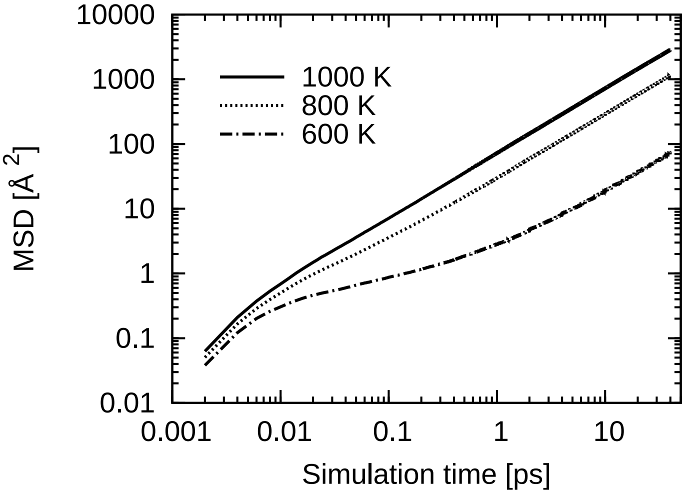

Figure 5.3: Mean square displacement of supercooled liquid GeTe as a function of time for selected temperatures.

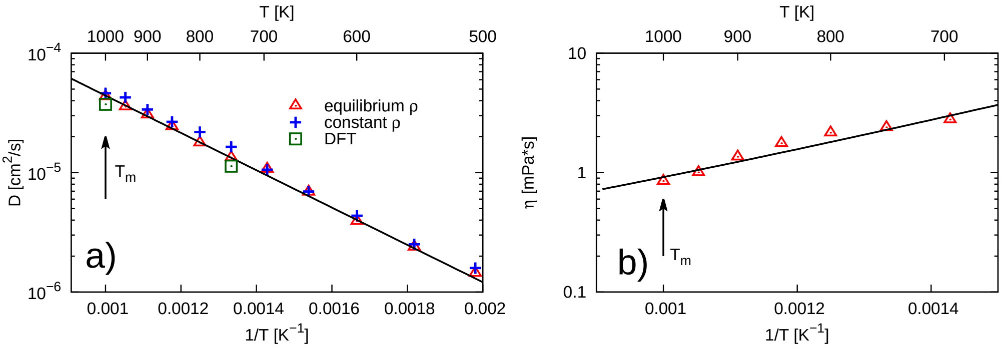

Figure 5.4: a) Self-diffusion coefficient D as a function of temperature in the supercooled liquid GeTe calculated from the mean square displacement (red triangles). The density is scaled with temperature according to the calculated thermal expansion coefficient. The open squares are the results of DFT simulations of a 216-atom cell. The crosses correspond to the values of D computed by holding the density fixed to the value at the melting point. b) Calculated viscosity η (Green-Kubo formula) as a function of temperature in the supercooled liquid. The density changes with temperature. The lines are Arrhenius fits of the data that give an activation energy of 0.220 ± 0.002 eV for D and 0.17 ± 0.035 eV for η. T_m is the theoretical melting temperature (see text).

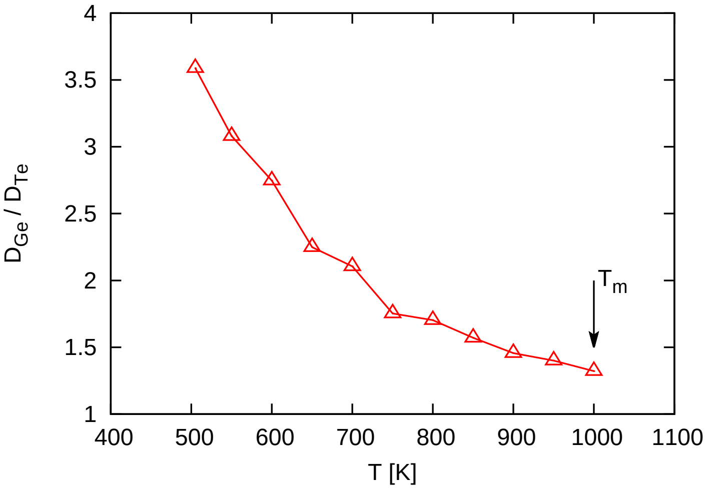

Figure 5.5: Ratio of the self-diffusion coefficients of the two species D_{Ge}/D_{Te} as a function of temperature in supercooled liquid GeTe.

be our crossover temperature by means of the Green-Kubo (GK) formula.

$$\eta = \frac{V}{k_B T} \int_0^{\infty} \langle \sigma_{xy}(t) \cdot \sigma_{xy}(0) \rangle dt \tag{5.3}$$

where $\sigma_{xy}(t)$ is the off diagonal component of the stress tensor. The time autocorrelation function of the stress tensor has been computed by means of the fast Fourier transform method [78] that exploits the convolution/correlation theorem by which the autocorrelation

$$\mathcal{C}_{\sigma\sigma}(t) = \langle \sigma_{xy}(t) \cdot \sigma_{xy}(0) \rangle \tag{5.4}$$

can be easily evaluated by taking the Fourier transform

$$\hat{\mathcal{C}}_{\sigma\sigma}(\nu) = \hat{\sigma}^*_{xy}(\nu) \cdot \hat{\sigma}_{xy}(\nu) \tag{5.5}$$

This product can easily take advantage of the very rapid algorithms available for computing Fourier transforms (FT), like FFTW [104]. Once computed, the inverse FT gives the autocorrelation. Care must be taken also in truncating the integral of Eq. 5.3. The stress autocorrelation function decays slowly, especially upon cooling, so that convergence of the integral with respect to truncation time must be carefully checked. The integral in Eq. 5.3 is converged by restricting the integration time to 60 ps due to the decay of the self-correlation function above the melting temperature. However, long simulation times up to 2 ns are needed to converge the average ($< .. >$) over different initial times $t = 0$.

Above T* the viscosity can be described by a simple Arrhenius (Fig. 5.4b) function with an activation energy of 0.17 ± 0.035 eV, very close to the value of 0.2 eV measured experimentally for the $Ge_{0.15}Te_{0.85}$ eutectic alloy above T_m[147]. For the GeTe composition, experimental values of η are available only at 1000 K yielding η=2.59 mPa·s which is twice as large as our result (cf. Fig. 5.4b). This discrepancy is not due to the NN potential but possibly to limitations of the underlying DFT framework. Previous works on $GeSe_2$ have indeed shown that different choices of the exchange and correlation functional affect the dynamical properties of the liquid phase [147]. The viscosity can be computed from the GK formula only above T^* since at lower temperatures the system crystallizes spontaneously on the time scale of few hundreds of ps which is not long enough to get the value of η converged. In the supercooled liquid, η can not be defined on a time scale longer than the crystallization time which in GeTe is very short in the temperature range 500-700 K.

We thus attempted to extrapolate η below T^* by a VTF-like function (see section 1.4) with the constraint of matching the typical value of 10^{15} mPa· s expected at T_g [16]. Unfortunately, a reliable value of T_g for is not available from experiments because of the fast crystallization of GeTe, and its theoretical estimate from simulations is uncertain as discussed at the end of this section. Experimental data on T_g are available for the better glass formers $Ge_xTe_{(1-x)}$ alloys with x=0.15-0.23 [148]. By a linear extrapolation of these latter data on T_g with x one obtains T_g=511 K for x=0.5 which is probably too high. On the other hand, T_g is customarily assumed to be slightly below the crystallization temperature, that is about 450 K in GeTe [149]. We then used the function proposed in [150] that allows fitting η over a wider range of temperatures

$$\begin{aligned} \log_{10}\eta(T) = \log_{10}\eta_o + (15-\log_{10}\eta_o)\cdot \\ \cdot\tfrac{T_g}{T}\exp\left[\left(\tfrac{m}{15-\log_{10}\eta_o}-1\right)\left(\tfrac{T_g}{T}-1\right)\right] \end{aligned} \tag{5.6}$$

where m and η_o are fitting parameters. In Eq. 5.6 η=10^{12} Pa· s at T_g.

The parameter m is is the fragility index of the supercooled liquid. The fragility of a supercooled liquid can be quantified by finding a measure of how much the viscosity deviates from the Arrhenius behavior with respect to the temperature expected in the case of a ideal strong liquid. In particular, the logarithmic derivative of η at T_g is defined by the fragility index

$$m = \frac{d(\log_{10}\eta)}{d(T_g/T)}|_{T=T_g} \tag{5.7}$$

We performed two fittings at two temperatures as shown in Fig. 5.6: the first with T_g=450 K that yields m=111 ($log_{10}\eta_o$=-0.18), and a second one with a somehow lower temperature T_g=400 K which yields a very similar value of m=104 ($log_{10}\eta_o$=-0.15). Similar results are obtained by using the modified VFT function proposed in Ref. [151]. Our data on viscosity above T^* are thus consistent with a high fragility of the supercooled liquid. For sake of comparison we remark that m=20 in silica which is a typical strong liquid while m=191 in PVC which is a typical fragile liquid [152]. Unfortunately, due to the lack of data on η below T* and the uncertainties in the value of T$_g$, we can not assign accurately the degree of fragility. Nevertheless, even for a value of m as large as 100, the viscosity rises too steeply in the range 500-600 K to be consistent with the calculated values of D and the application of the Stokes-Einstein relation. In the hydrodynamic regime when the SER holds it is actually possible to estimate the viscosity on the shorter time scale of 50 ps by a finite size scaling analysis [153] of the self-diffusion coefficient based onto the SER:

$$D(L) = D_\infty - \frac{2.387 k_B T}{6\pi\eta L} \tag{5.8}$$

where D_∞ is the non size dependent diffusion coefficient, and L is the edge of the cubic simulation cell. It is worthwhile to remember that Eq. 5.8 can be applied only in the hydrodynamic regime when the SER (Eq. 1.2) holds. One can perform different MD simulations at the same temperature with different system sizes (different L) and obtain η from Eq. 5.8.

We considered three cubic models with 512, 1728 or 4096 atoms. By applying Eq. 5.8 above T*, we obtained values for η very close to the GK data (Fig.5.7) and consistent with the SER. However, when Eq. 5.8 is applied below T*, one obtains values of η that are three orders of magnitude larger than those obtained from D and the application of the SER ($\eta = \frac{k_B T}{6\pi R D}$, where R is the average van der Waals radius of the two species) as shown in Fig.5.7. This inconsistency demonstrates that the SER indeed breaks down. We remark that the numerical values of η reported in Fig.5.7 below T* are not reliable since they are obtained from Eq.5.8 which is not applicable when the SER breaks down. Note also that the values of η obtained from the scaling of D above T* are better fitted by an Arrhenius function than the GK data (cf. Fig.5.4b) possibly because of numerical inaccuracy in the GK values at the lowest temperatures.

5.3 Overheated amorphous phase

We have seen how atomic mobility in supercooled liquid GeTe remains high even in the neighborhood of the glass transition temperature. Now we can ask ourselves if the amorphous phase would display the same properties once heated above the glass transition. Thus, we investigated the overheated amorphous phase of GeTe applying exactly the same framework used in the previous section for the supercooled liquid. We chose a density of 0.0334 atoms/Å^3 [134] very close to the experimental one for the amorphous phase [129]. We used the two models generated by quenching from 1150 K to 300 K in 100 or 300 ps which yielded the same structural and dynamical properties [134]. The volume was then increased according to the calculated linear thermal expansion coefficient of α=11$\cdot$10^{-6} K^{-1} at 300 K which is in the range of values measured for other materials in this class [146]. We increased the temperature from 300 K to 700 K in 5 steps. At each step the system is equilibrated for 25 ps with the thermostat. Overall the system is thus heated

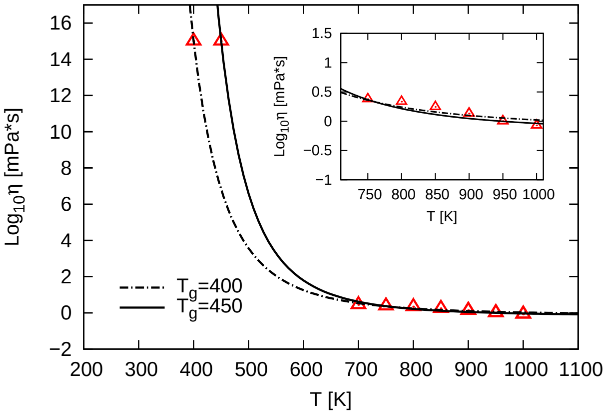

Figure 5.6: Viscosity as a function of temperature in supercooled liquid GeTe. The lines are fittings with Eq. 5.6 of the Green-Kubo data (triangles) above T* under the constraint of reproducing the value of η=10^{12} Pa·s at T$_g$ with T$_g$=450 K (continuous line) or T$_g$=400 K (dashed line, see text). A magnification of the data for T>T* is shown in the inset.

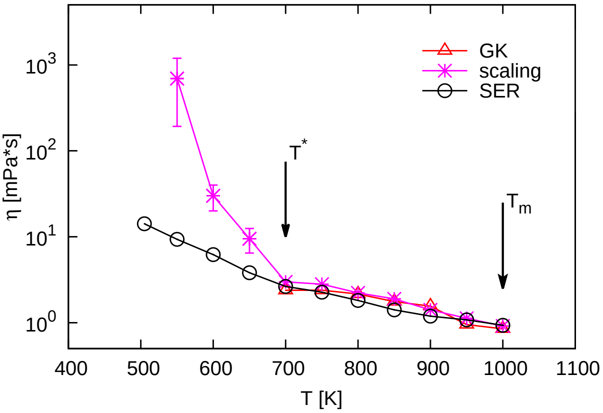

Figure 5.7: Viscosity computed from the Green-Kubo (GK) formula (red triangles), from the scaling of the diffusion coefficient with the simulation cell size (stars), and from the self diffusion coefficient of the 4096-atom cell and the application of the SER (open circles).

from 300 K to 700 K in 125 ps. At each temperature statistical averages are then collected on longer microcanonical simulations.

We first computed D from the atomic mean square displacement on the time scale of 50 ps. The values of D as a function of temperature are reported in Fig. 5.8 and compared with the corresponding values in the supercooled liquid phase. The ratio between the self-diffusion coefficient of Ge and Te (D_{Ge}/D_{Te}) decreases by increasing temperature as shown in Fig. 5.9. We remark that the temperature in the supercooled liquid is changed in time along the curve in Fig. 5.8 following a protocol similar to that used to generate the amorphous model. In fact the liquid was quenched from 1000 K to 500 K in 250 ps along the curve in Fig. 5.8 while the amorphous models were generated by quenching from 1150 K to 300 K in either 100 ps or 300 ps with the same results.

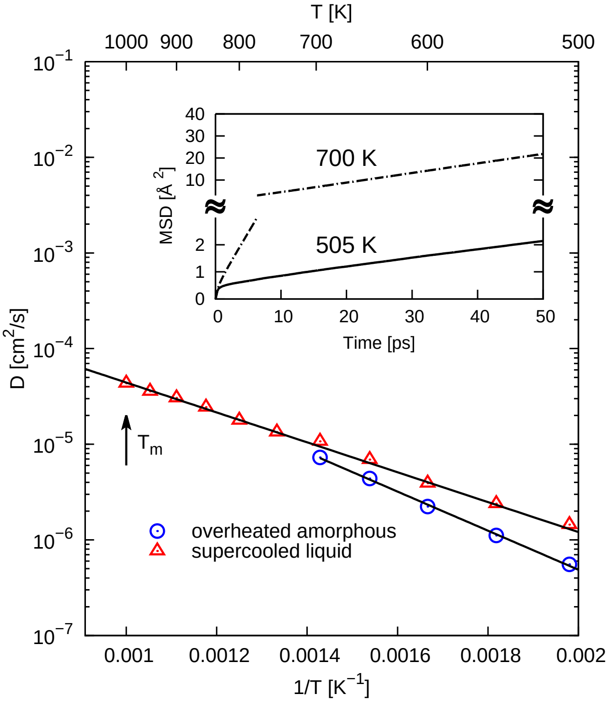

Figure 5.8: Self-diffusion coefficient D as a function of temperature calculated from the mean square displacement in the overheated amorphous phase (open circles) and in the supercooled liquid phase. The straight lines are Arrhenius fits of the data that give an activation energy of 0.405 ± 0.008 eV for the overheated amorphous phase and 0.220 ± 0.002 eV for the supercooled liquid. T_m is the theoretical melting temperatures. Inset: Atomic mean square displacement as a function of time in the overheated amorphous phase at 505 K and 700 K.

Clearly a hysteresis is present, the values of D are lower in the amorphous phase than

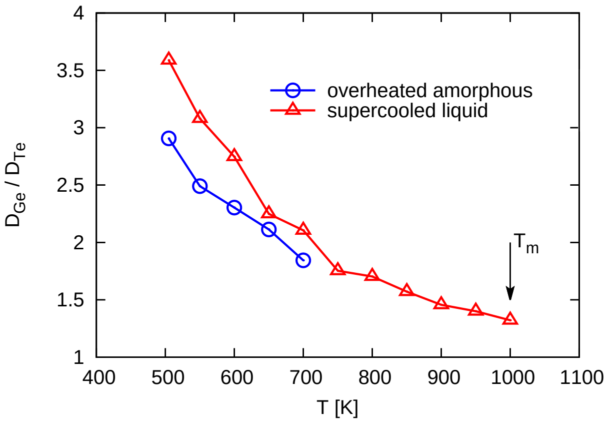

Figure 5.9: Ratio of the self-diffusion coefficients of the two species D_{Ge}/D_{Te} as a function of temperature in the overheated amorphous GeTe (open circles) and in the supercooled liquid phase.

in the supercooled liquid it originates from. The overheated amorphous is in fact more structured than the supercooled liquid at 505 K as shown by the comparison of their pair distribution functions (PDF) in Fig. 5.10a. At 700 K the differences in the PDF of the overheated amorphous and the supercooled liquid are instead very small (cf. Fig. 5.10b). Whether it is necessary or not to distinguish between the two phases in the phase change memory cell and to take into account the hysteresis in the glass transition, it depends on the details of the programming current in the set operation which controls the value of the intermediate temperature between T_g and T_m reached during the recrystallization process.

To assess the validity of the SER, the viscosity in the overheated amorphous phase as a function of temperature should be computed. However, a direct calculation of the viscosity is possible only for temperatures at which the system, in the supercooled liquid or in the amorphous phases, does not crystallize on the time scale needed to define/compute the viscosity. The Green-Kubo formula (see Eq. 5.3) requires simulation time about 1 ns long to converge the average ($< .. >$) over initial states. On this time scale the overheated amorphous models spontaneously crystallize for temperatures below 700 K, preventing a reliable estimate of the viscosity exactly as observed in the supercooled liquid. Nevertheless, it is possible also in the case of overheated amorphous GeTe to assess the validity of SER by studying the finite size scaling of the self-diffusion coefficient on the shorter time scale of 50 ps. In the hydrodynamic regime when the SER holds it is actually possible to estimate the viscosity from the scaling of D with the edge L of the cubic simulation cell as Eq. 5.8. We again considered three models with 512, 1728 and 4096 atoms at the same atomic density. By applying Eq. 5.8 to the supercooled liquid above 700 K, where the system does not crystallize on the time scale of about 1 ns, we obtained values for η very

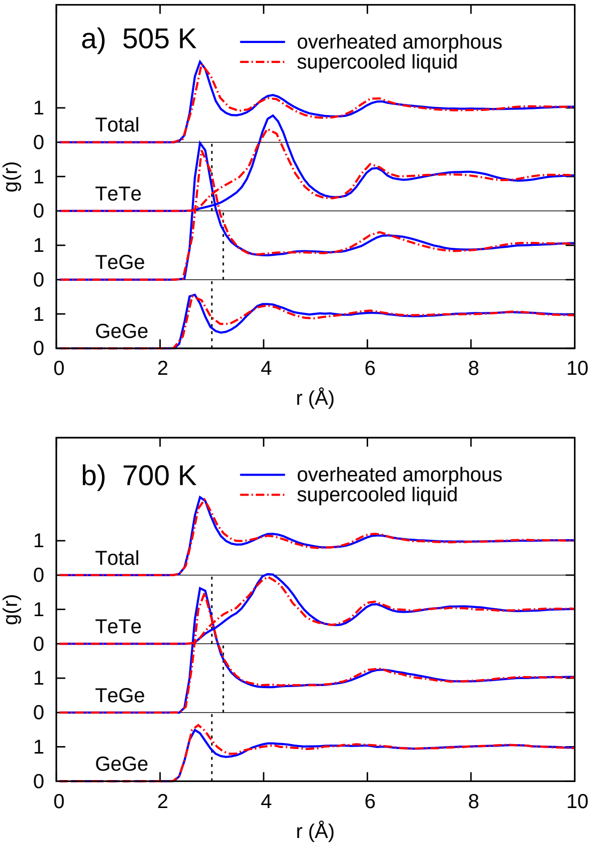

Figure 5.10: Total and partial pair correlation function of the supercooled liquid and the overheated amorphous GeTe at a) 505 K b) 700 K.

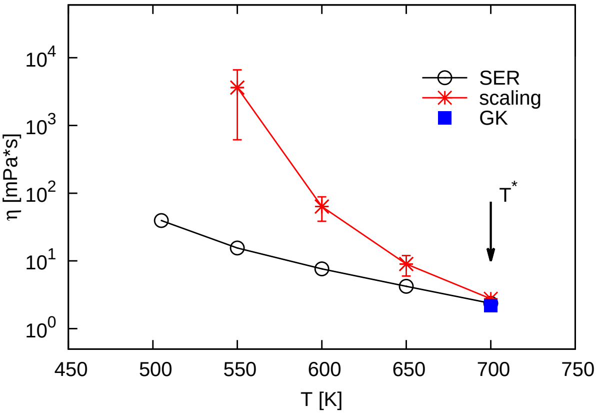

Figure 5.11: Viscosity computed from the scaling of the diffusion coefficient with the simulation cell size (stars), and from the self diffusion coefficient of the 4096-atom cell and the application of the SER (open circles) in the overheated amorphous phase. The square at 700 K is obtained from the GK formula.

close to that computed directly from the Green-Kubo formula (Fig. 5.11) and consistent with the SER. However, when Eq. 5.8 is applied below 700 K in supercooled liquid and here in the overheated amorphous phase one obtains values of η that are three orders of magnitude larger than those obtained from D and the application of the SER ($\eta = \frac{k_B T}{6\pi R D}$, where R is the average van der Waals radius of the two species). This inconsistency, shown in Fig. 5.11, demonstrates that the SER indeed breaks down in the overheated amorphous phase as well as in the supercooled liquid below 700 K. We remark that the numerical values of η reported in Fig. 5.11 below 700 K are not reliable since they are obtained from Eq. 5.8 which is not applicable when the SER breaks down (Fig. 5.11). At 700 K the two approaches yielded the same result which is also equal to the value of η computed directly from the GK formula. Note that η from the GK formula at 700 K is 2.2 mPa·s in the overheated amorphous phase and 2.4 mPa·s in the supercooled liquid.

As discussed previously. a reliable experimental value of T_g is not available for GeTe nor for the other fast crystallizing phase change compounds. In principle T_g could be estimated for the temperature dependence of the volume. In fact, by freezing the liquid sufficiently fast to prevent crystallization we would observe a decrease in the specific volume V as sketched in Fig. 5.12 down to the glass transition temperature T_g at which a change in the slope of $V = V(T)$ is expected.

On the contrary the specific volume experiences a discontinuous jump upon crystallization at T_m (see Fig. 5.12). The glass transition temperature T_g of GeTe could then be estimated by a plot of the type sketched in Fig. 5.12 for our system. Unfortunately the equilibrium volume of the liquid phase of GeTe strongly depends on the choice of the vdW correction to the NN potential that is needed to reproduce the equilibrium density

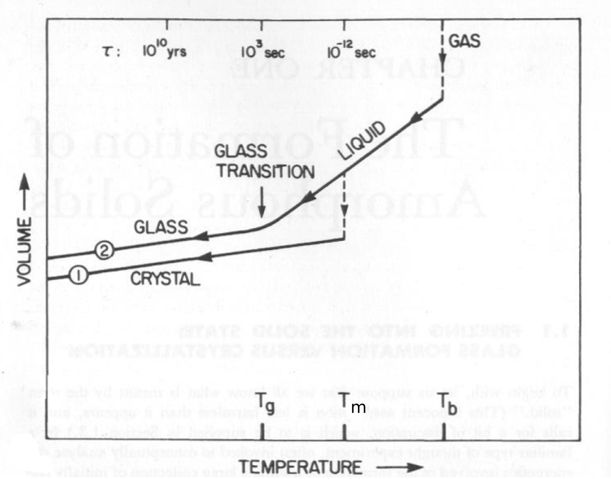

Figure 5.12: The two general cooling path by which a system can condense into the solid phase. Route ① is the path to the crystalline state; route ② is the rapid-quench path to the amorphous solid state. From Ref. 154.

of the liquid as discussed in section 2.3.4. Nevertheless we can attempt to estimate T_g by assuming a vdW correction equal for the amorphous and the liquid phase obtained by matching the experimental equilibrium density of the liquid at T_m. The resulting volume versus temperature points for our models of the amorphous and liquid phases of GeTe are reported in Fig. 5.13, which yields T_g=505 K.

Although this result looks like a reasonable value, we remark that the equilibrium volume of the amorphous at room temperature is now 10% smaller then in experiments, which suggests that the vdW correction could not be taken equal for the two phases. As a final remark we mention that the fast crystallization of GeTe and GeSbTe alloys has been ascribed to the similarity of the bonding topology in the amorphous and crystalline phases [17]. In fact, the must abundant rings in amorphous GeSbTe and GeTe are the four-membered ABAB rings (A=Ge/Sb and B=Te) which are also the building blocks of the cubic crystalline phase. The presence of nanocavities [8] in the amorphous phase has also been proposed as a structural feature that aids the alignment of the four-membered rings during crystallization. The results presented here suggest that in addition to the presence of these structural features in the amorphous phase, the crystallization is promoted also by a high atomic mobility just above T_g where, due to high supercooling, a large driving force ($\Delta\mu$) boosts the crystallization speed. A large atomic diffusivity coexists with a large viscosity due to the breakdown of SER both in the supercooled liquid and in the overheated amorphous phases. Therefore the self-diffusion coefficient to be used in the modeling of the crystallization process by classical nucleation theory can not be inferred from the expected viscosity and the SER, nor the crystallization speed can be extrapolated from measurements below T_g. Indeed the crystallization of the amorphous phase at low

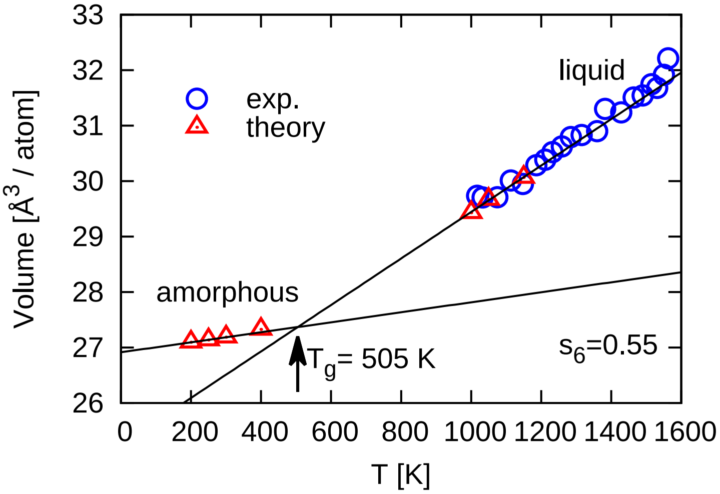

Figure 5.13: Volume versus temperature plot for amorphous and liquid GeTe. Each point has been obtained from the equation of state at finite temperature. We underline that the value of T_g obtained is not reliable at all, while thermal expansion coefficients -coming from the slopes of the curves- are correct due to the fact that they are not affected by Van der Waals corrections. Experimental points was adapted from Ref.116.

temperatures below T_g of interest for data retention might take place in a different manner with respect to the crystallization of the highly mobile overheated amorphous phase and of the supercooled liquid above T_g. Large scale simulations based on the NN potential hold the promise to shed light on the crystallization at low temperatures below T_g as well in the near future. The results discussed in this chapter on the SL have been published in Ref. [155].

Chapter 6

Crystallization of Amorphous and Liquid GeTe

6.1 Homogeneous crystallization

We investigated homogeneous nucleation and growth by direct molecular dynamics simulations. We considered the supercooled liquid (SL) GeTe at 550, 600, 650 and 700 K. Models of 512, 1728 and 4096-atoms have been considered to study finite size effects. We also performed simulations at the same temperatures of the overheated amorphous (OA) GeTe. The density of each phase was scaled according to the calculated thermal expansion coefficients (see Fig. 5.13), and simulations lasted 1 ns in the NVT ensemble after 25 ps of equilibration at each temperature. In Fig. 6.2 we show the potential energy profiles for both the SL and the OA. The OA does not crystallize at all at 700 K, due to the fact that approaching T_m the nucleation rate decreases. We actually observed crystallization at 700 K for SL, in which by chance a single critical nucleus pops up and immediately grows very fast.

In order to measure the crystal growth speed, we have to identify the atoms belonging to the crystalline nuclei. To this aim, we used a local order parameter constructed according to Steinhardt [156] and Frenkel [157] as:

$$\eta(i) = \sum_{j=1}^{N_b(i)} \frac{\sum_{m=-l}^{l} q_{6m}(i) q_{6m}^*(j)}{\sqrt{\sum_{m=-l}^{l} |q_{6m}(i)|^2} \sqrt{\sum_{m=-l}^{l} |q_{6m}(j)|^2}} \tag{6.1}$$

where

$$q_l(i) = \sqrt{\left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2 \right)} \tag{6.2}$$

and

$$q_{lm}(i) = \frac{1}{\sum_{j=1}^{N_b(i)}} \sum_{j=1}^{N_b(i)} Y_{lm}(\vec{r}_{ij}) \tag{6.3}$$

where $N_b(i)$ is the number of bonds for particle i defined within a cut-off radius of 3.2 Å, Y_{lm} are spherical harmonics and $\vec{r}_{ij}$ are interatomic distances vectors. The $\eta(i)$ order parameter is able to discriminate with great accuracy between the amorphous and the crystalline phase as shown in Fig. 6.1.

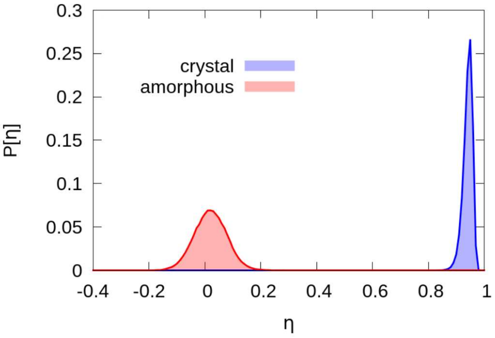

Figure 6.1: Order parameter η distribution for crystalline and amorphous GeTe

We identify the size of the critical nucleus as the minimum size of the cluster that one formed dose not disappear but continue growing until the model is fully crystallized. A more compelling definition requires the calculation of the nucleus free energy as a function of time which is beyond the scope of this work

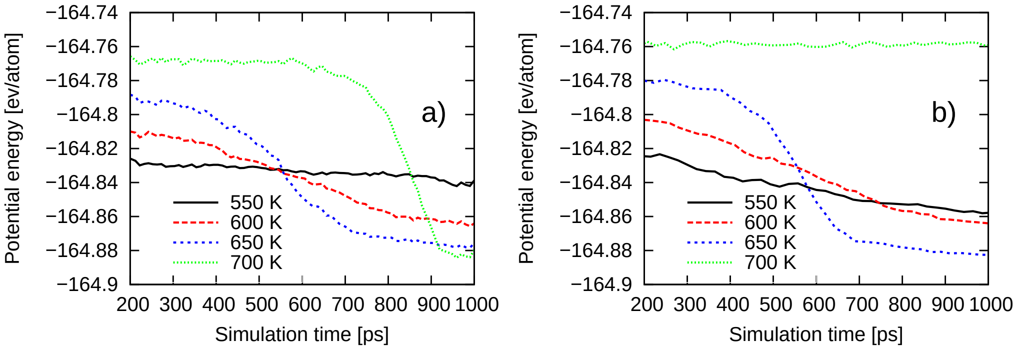

Figure 6.2: Potential energy of SL a) and OA B) GeTe as a function of simulation time at different temperatures.

In Fig. 6.3 we report the temperature dependence of the critical nucleus size. When more than one column is drawn for a certain temperature, it means that for that temperature more than one critical nucleus size is present at the same time and grows together with the other nuclei. As expected from Classical Nucleation Theory (CNT), the size of the critical nucleus increases with temperature.

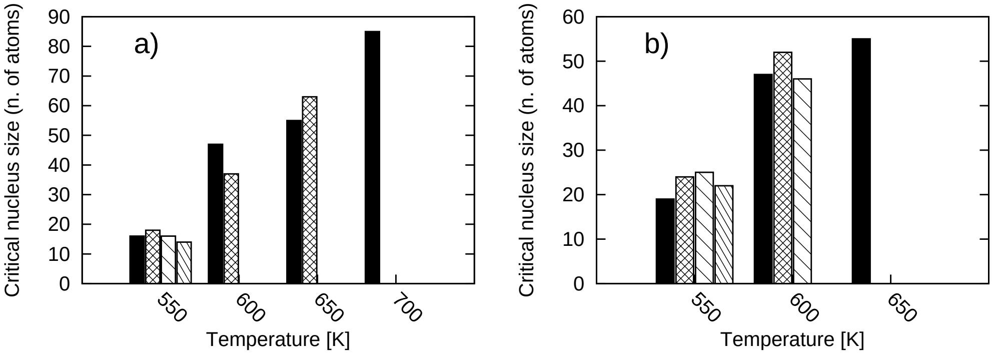

Figure 6.3: Temperature dependence of critical nucleus size for a) Supercooled liquid and b) overheated amorphous GeTe.

By using the order parameter defined in Eq. 6.1, we tracked the number of crystalline atoms in the simulations. The results are shown in Fig. 6.4 for both SL and OA. The crystallization speeds appear similar in the two phases, although the OA crystallizes faster than SL at 550 K and slower than SL at 600 and 650 K. In principle, SL should crystallize faster then OA because the diffusion coefficient of the former is larger than that of the latter, and the kinetic prefactor $U_{kin}(T)$ of Eq. 1.7 depends on the diffusion coefficient. Still, this effect is competitive with the fact that the free energy difference between the amorphous and the crystal is different from the free energy difference between the liquid and the crystal, so that it is possible that crystallization speed is faster for OA in certain temperature ranges only. At low temperatures the nucleation probability is high. In fact, as shown in Fig. 6.3, at low temperatures there are several supercritical nuclei.

On the other hand, at high temperatures, e.g. SL at 700 K, nucleation probability is low. In that particular simulation, only one nucleus appears and grows (see Fig. 6.5). After an initial transient, the number of crystalline atoms grows in a perfect cubic fashion until eventually the nucleus starts interacting with its periodic images (this is the region in Fig. 6.4 in which the growth curve becomes a constant), leading to an almost complete crystallization of the model. The fraction of uncrystallized material depends on temperature.

Since different nuclei can coexist and grow at the same time, different crystalline grains are formed with different crystallographic orientations, as shown in Fig. 6.6.

An important issue in simulating crystallization is the effect of periodic boundary conditions. Because of the interaction with periodic images, in small cells (512-atom) it was actually difficult to estimate a speed of crystal growth since the model was fully

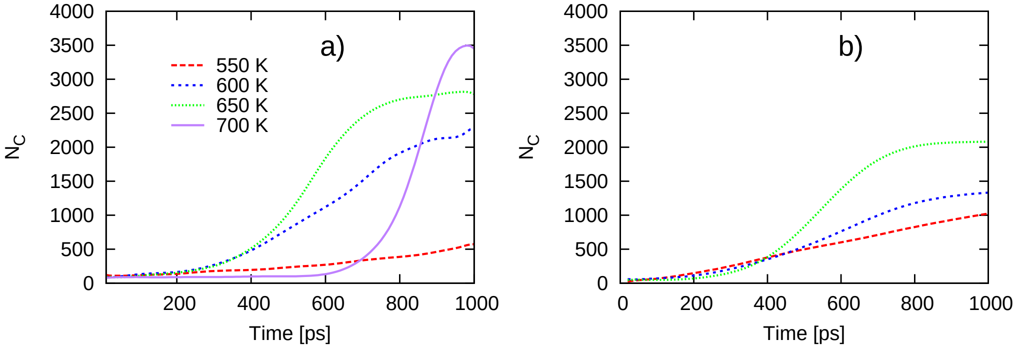

Figure 6.4: Fraction of crystalline atoms (N_c) as a function of time for a) supercooled liquid and b) overheated amorphous GeTe.

crystallized before the growth of the nucleus reached a steady state. The 4096-atom models are actually large enough to provide an estimate of the growth speed. In Fig. 6.7 we show a typical example. In panel a), the crystalline nucleus interact with its periodic images; in panel b) a nucleus of the same size can grow freely in a larger 4096-atom cell. In panel c) and d) an example of an almost spherical and of a strongly elongated critical nucleus are shown.

CNT assumes that nuclei are spherical. This is not true in general, and many attempt to overcame this assumption can be found in literature (see for example Ref. 45). In GeTe, we observed that the higher the temperature, the more the nuclei tend to become spherical. We quantify the deviation of a nucleus shape from a perfect sphere by the asphericty parameter [158]

$$s_0 = \frac{(I_{xx} - I_{yy})^2 + (I_{xx} - I_{zz})^2 + (I_{yy} - I_{zz})^2}{2(I_{xx} + I_{yy} + I_{zz})^2} \tag{6.4}$$

where I_{ij} are the elements of the tensor of inertia defined by

$$I_{ij} = \sum_{k=1}^{N} m_k (r_k^2 \delta_{ij} - r_{ki} r_{kj}) \tag{6.5}$$

where r_{ki} is the i-th component of the vector between the center of mass and particle k, r_k is the overall magnitude of this vector and δ_{ij} is the Kronecker delta. The eigenvalues of this tensor are the principal components of the moment of inertia used in Eq .6.4 to determine s_0, which ranges from zero for a perfectly spherical cluster unity for extremely elongated ones. The nuclei with $s_0 < 0.2$ appears very much spherical by visual inspection. We remark that while s_0 measure spherical symmetry, it does not measure the compactness. Asphericity of the nuclei is shown in Fig. 6.8.

In order to obtain the crystal growth speed, we calculate the volume of each nucleus

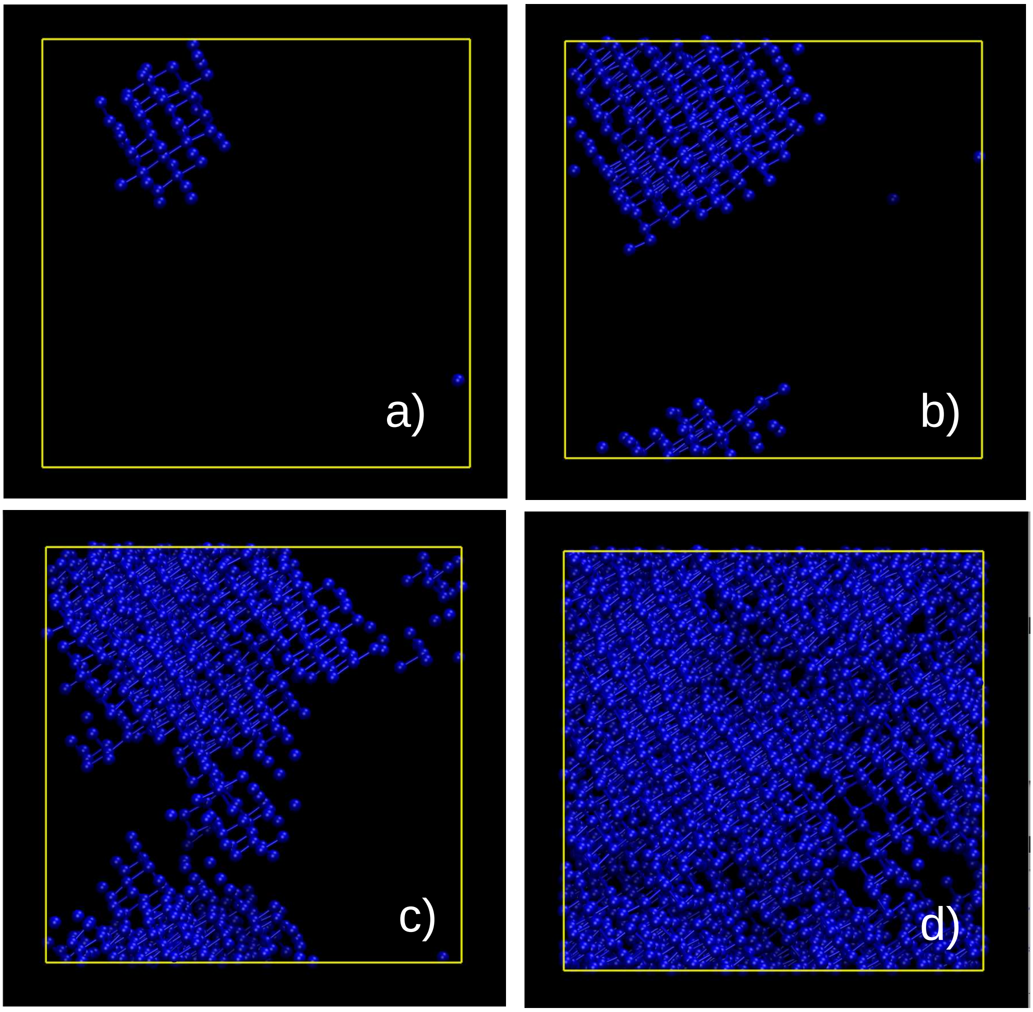

Figure 6.5: Crystallization of a supercooled liquid model of 4096-atom at 700 K. Panel a) shows the critical nucleus, that rapidly (snapshots are taken every 20 ps) grows (panels b) e c)) until almost the whole system is crystallized (panel d)) in a single grain.

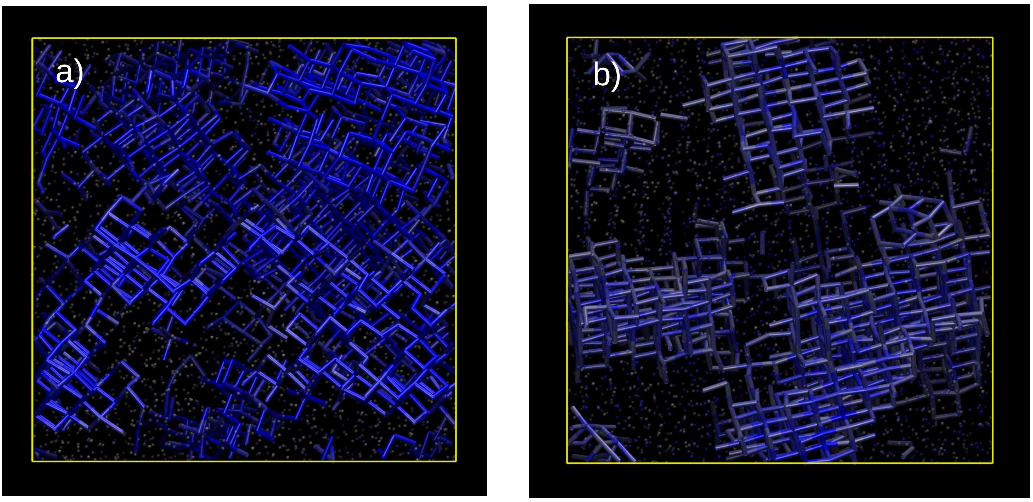

Figure 6.6: Different crystalline grains. a) Snapshot of a MD simulation of a 4096-atom OA at 600 K after 1 ns. Different nuclei are present, each one displaying a different orientation. b) Snapshot of a MD simulation of a 4096-atom SL at 550 K after 1 ns, with a single grain interacting with its periodic images.

as a function of time. One can then approximate the nucleus with a sphere (or a cube), and obtain the dependence of the radius (or the edge) on the simulation time. This gives a rough estimate of the linear crystal growth speed measured in experiments. In particular, we chose to define the radius as

$$R(t) = \frac{1}{2}\sqrt[3]{V_{cr}(t)} \tag{6.6}$$

where the volume of a crystalline nucleus V_{cr} is computed as the sum of the volumes of the Voronoi polyedra centered on the atoms of the nucleus identified by the order parameter defined in Eq. 6.1. To this aim we used a modified version of the Voro++ library [159]. In Fig. 6.9 we show the radius of the growing nucleus as a function of time for selected temperatures. When more than one nucleus is present (see Fig. 6.9) the crystal growth speed has been calculated as the average of the speeds of the different nuclei.

In Fig. 6.10 the crystal growth speed of SL and OA are shown, together with the experimental data for GST from the work of Orava [53]. Since the diffusion coefficient D and the free energy difference between the solid and the liquid is different from SL and OA, the two phases display different crystallization speeds.

In order to obtain the crystal growth speed $U(T)$ at all intermediate temperatures, we calculated the free energy difference $\Delta\mu(T)$ according to Eq. 1.9, using our estimate for the melting temperature and the enthalpy of fusion. The latter was obtained from the enthalpy difference between the liquid and the crystal at their equilibrium volumes at the melting temperature, ΔH_m =36.29 kJ/mol. Using Eq. 1.7 we can then extract the

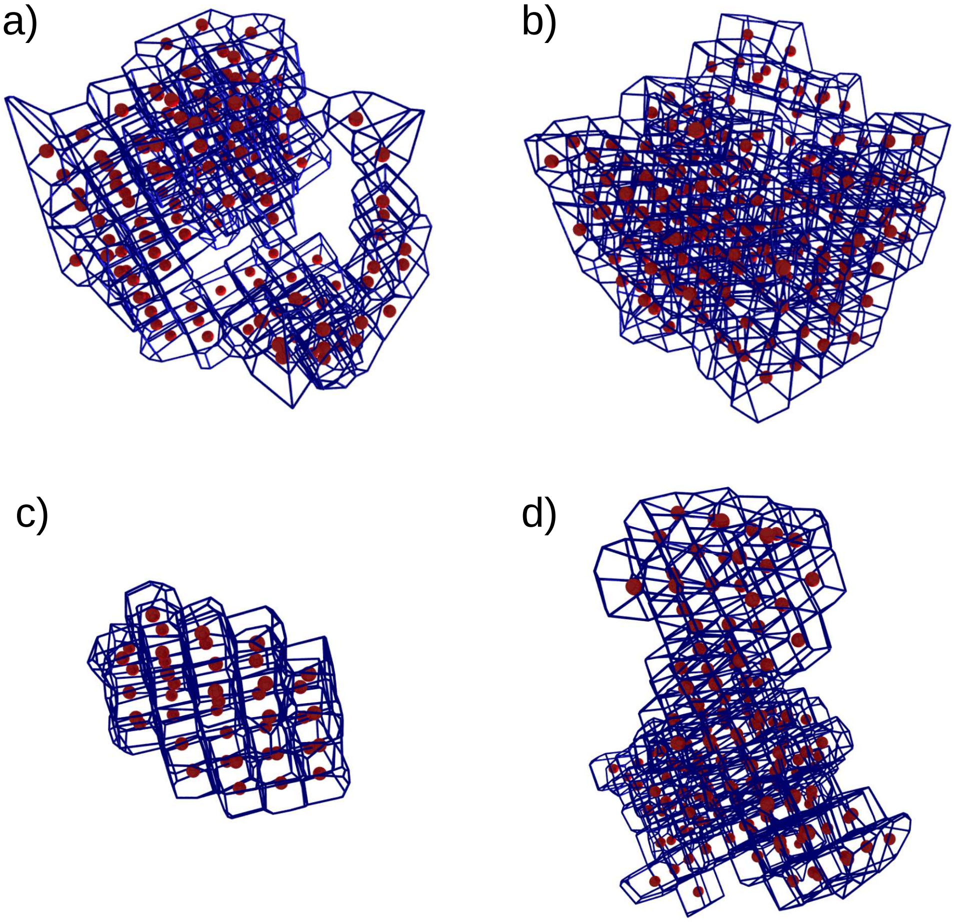

Figure 6.7: Asphericity of crystalline nuclei. In panel a) the effect of periodic boundary conditions on the crystalline growth in a small (512-atom) simulation cell. In panel b), a growing nucleus containing more or less the same number of atoms displayed in panel b). In this case, the simulation cell contains 4096 atoms, an the nucleus can grow freely without interacting with its periodic replica. Both simulations were performed at 600 K with models of supercooled liquid GeTe. In panel c) an almost spherical critical nucleus in amorphous GeTe overheated at 550 K. In panel d) a strongly non spherical nucleus growing in amorphous GeTe overheated at 650 K. Both simulations were performed with 4096-atom models.

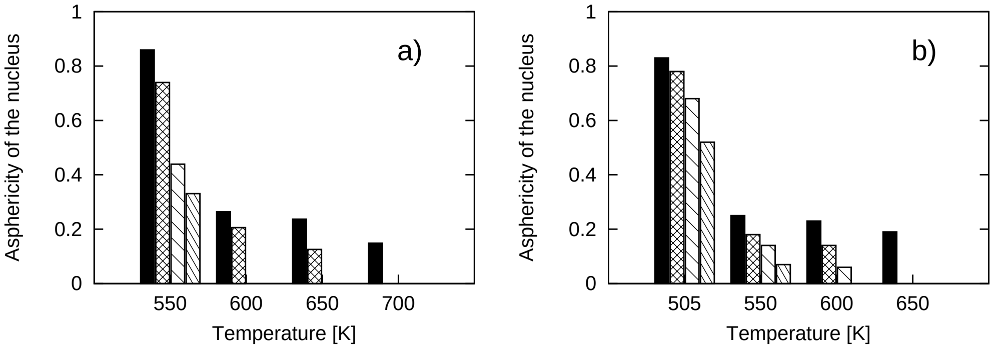

Figure 6.8: Asphericity of the critical nuclei as defined by Eq. 6.4, as a function of temperature. Each column refer to a single nucleus, e.g. at 550 K we have four nuclei, while at 600K we have two. a) SL, b) OA.

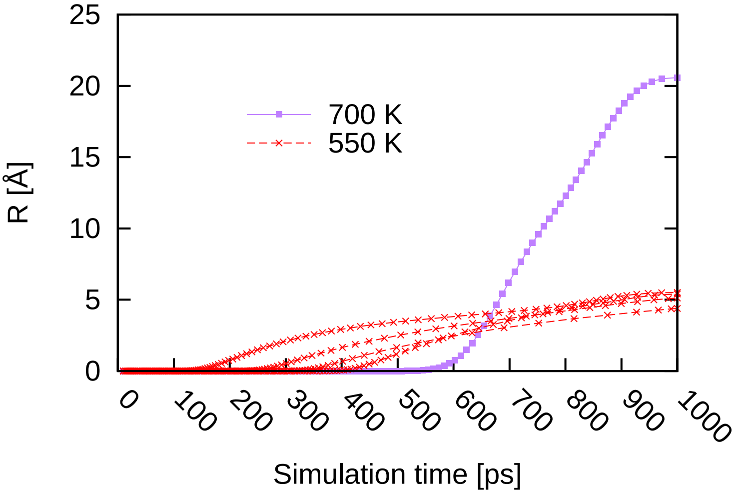

Figure 6.9: Radius of the growing nucleus for SL at two different temperatures, calculated according to Eq. 6.6. At 700 K a single nucleus grows, while the four lines at 550 K refer to four different nuclei.

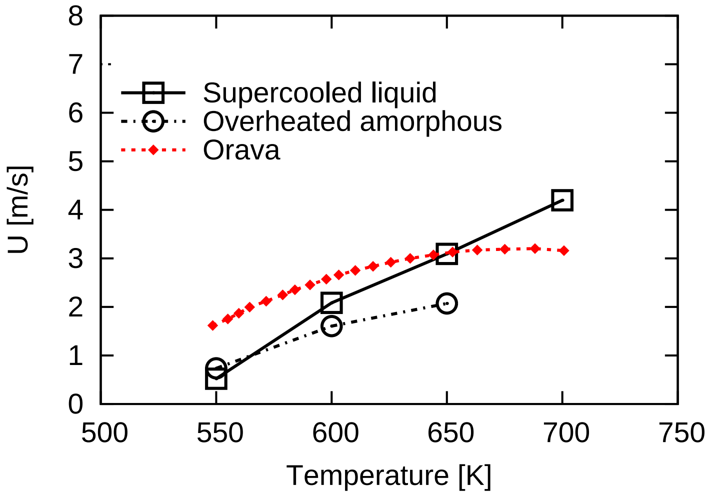

Figure 6.10: Crystal growth speed as a function of temperature for homogeneous crystallization of supercooled liquid and overheated amorphous GeTe. Data refer to 4096-atom models. The experimental data for GST obtained from differential scanning calorimetry by Orava *et al.* [53] are also shown.

kinetic prefactor $U_{kin}(T)$ from $U(T)$ for the simulated temperatures (550-700 K). The results are shown in Fig. 6.11 for SL and also for OA although it has been assumed that $\Delta H_m(liquid - crystal) = \Delta H_m(liquid - amorphous)$.

By fitting $U_{kin}(T)$ with an Arrhenius equation we obtained activation energies for crystallization of 0.26 and 0.22 eV for SL and OA respectively. These values are very similar to the activation energies extracted from the temperature dependence of the diffusion coefficients of SL and OA reported in section 5.3 and section 5.2. U_{kin} is 5.6·10^2 cm/s for SL at 700 K (see Fig. 6.11). If we substitute this value in Eq. 1.8 with $\lambda = 3$ Å and using the computed value of the diffusion coefficient at 700 K, $D = 1.1 \cdot 10^{-5}$ cm^2/s, we obtain a reasonable value of $\gamma_s = 0.25$ which suggests that the high crystallization speed is indeed due to the large atomic mobility. Arrhenius fits define the temperature dependence of U_{kin}, that we can now insert in Eq. 1.7 obtaining the crystal growth rate in the whole temperature range (T_g-T_m). The result is shown in Fig. 6.12. The maximum of $U(T)$ is shifted towards higher temperature for SL and OA GeTe with respect to what has been observed for GST. This is due to the fact that GeTe has a melting temperature about 100 K higher than GST. These plots are less reliable for OA because of the uncertainties in ΔH_m for this phase. Actually ΔH_m should be higher for the amorphous-crystal transition than for the liquid-crystal transition. The value of U_{kin} of OA higher than that of SL at 550 K in Fig. 6.11 is possibly due to the approximation of setting ΔH_m equal for the amorphous and the liquid, which becomes less accurate by decreasing temperature.

A more detailed understanding of the crystallization mechanism in phase change mem-

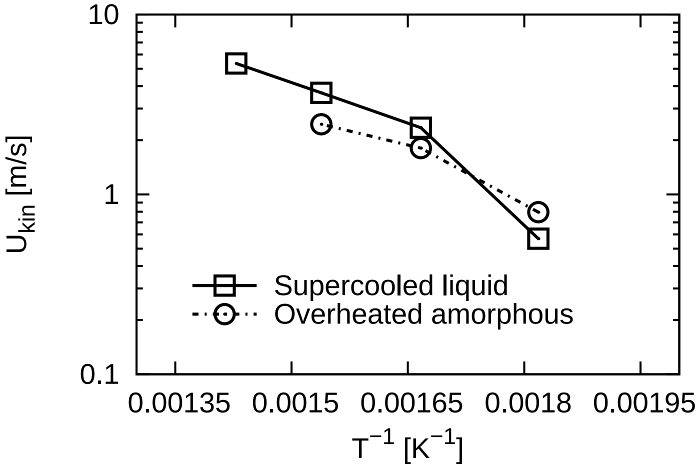

Figure 6.11: Kinetic coefficient for crystal growth (Eq. 1.7) from homogeneous nucleation of supercooled liquid and overheated amorphous GeTe. Data refer to 4096-atom models.

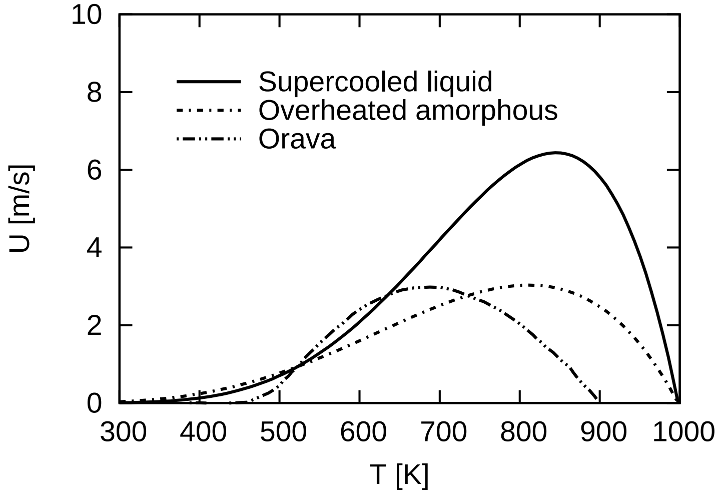

Figure 6.12: Crystal growth speed as a function of temperature for homogeneous crystallization of supercooled liquid and overheated amorphous GeTe. obtained by assuming $U(T)$ and U_{kin} in the form of Eq. 1.7 and Eq. 1.8 with parameters obtained from the simulations at selected temperatures (see text).

ory cells has to take into account heterogeneous crystallization at the boundary between the amorphous phase and the surrounding crystalline matrix. Simulations of the crystallization at the liquid/amorphous- crystal interfaces are ongoing.

Conclusions

We developed an interatomic potential for the bulk phases of GeTe by means of a novel approach based on Neural Networks algorithms, which allowed us to investigate several properties of the phase change material GeTe that are currently beyond the reach of DFT atomistic simulations. Starting from a huge dataset of DFT energies, we construct an analytical interatomic potential that retains a quasi ab initio accuracy and allows us to perform molecular dynamics simulations of thousands of atoms for several nanoseconds, overcoming the limitations of DFT molecular dynamics in terms of system size and simulation timescale.

The NN potential has been validated by comparing the results on structural and dynamical properties of the bulk phases of GeTe with previous data from DFT calculations. By means of NN simulations, we have been able to study the dependence of the structural properties of a-GeTe on quenching time and system size. Actually, no sizable changes are observed in the structural properties of the large 4096-atom cell by increasing the quenching time from 100 to 300 ps from the melting temperature to 300 K. On the other hand, we observed sizable fluctuations in structural properties among 10 independent 216-atom models of a-GeTe all quenched in 100 ps both in NN and DFT simulations. However, averaging over 10 models is enough to obtain results very close to those of the large 4096-atom cell. The NN potential developed here for the GeTe stoichiometric composition can be used also for $\mathrm{Ge}_x\mathrm{Te}_{1+x}$ alloys with small x. However, the potential is not yet transferable to strongly off-stoichiometric composition such as the eutectic alloy $\mathrm{Ge}_{0.15}\mathrm{Te}_{0.85}$ because it is unable to reproduce accurately the interaction between long Te-Te chains as we explicitly checked by means of DFT simulations (no Te-Te chains but few dimers are present in $\mathrm{Ge}_{0.5}\mathrm{Te}_{0.5}$). The transferability of the NN potential can, however, be systematically extended to the whole binary phase diagram of the $\mathrm{Ge}_x\mathrm{Te}_{1+x}$ alloy by including new DFT configurations with different compositions in the database. The simulations with the NN potential are still sizably more expensive than simulations with classical force fields but does provide a huge speed up with respect to DFT simulations. The computational load is 1.5 min/ps on a 256 cores of Cray XT5 for the 4096-atom cell, which is five to six orders of magnitude faster than a standard Born-Oppheneimer DFT simulation with the CP2K code or at least four orders of magnitude faster than simulations with wave-function extrapolations either with CP2K or with the plane-wave

CPMD code. Thus, the development of a classical potential with close to DFT accuracy represents a breakthrough in the simulation of phase change materials, as it will allow addressing several key issues on the properties of this class of materials that are presently beyond the reach of DFT simulations.

As a first application of our NNP, we investigated thermal transport in the amorphous phase of the GeTe compound. We computed the thermal conductivity from equilibrium molecular dynamics and the Green-Kubo formula. The resulting value κ= 0.27 ± 0.05 W/(mK) is within the range of values measured for several other GeSbTe phase change materials. The calculation of phonon group velocity and phonon lifetimes reveals that essentially all phonons have a short mean free path in the range 1-10Å. The contribution to the thermal conductivity by propagating modes is thus negligible. The thermal conductivity is in fact mostly due to nonpropagating, delocalized modes (diffusons) which can transport heat according to the theory developed by Allen and Feldman [40]. Our results thus suggest that the bulk value of κ measured for instance in thin films can indeed be used to model the thermal transport of GeTe also in nanoscale devices, as the contribution from propagating modes that may endure ballistic transport at the scale of 10-20 nm is negligible. Phase change ternary alloys GeSbTe have lattice thermal conductivity very similar to that found here for GeTe. The phonon spectrum and the structure of amorphous GeTe are also very similar to those of other GeSbTe compounds with coexistence of tetrahedra and defective octahedra and the presence of nanovoids. It is therefore conceivable that the absence of phonons with long mean free path and the dominance of nonpropagating, delocalized modes in the thermal transport found here for GeTe might be extended to other materials in the same class.

We then focused on the dynamical properties of the supercooled phase of GeTe in the attempt to shed light onto the microscopic origin of its fast crystallization. We have demonstrated by means of MD simulations that the supercooled liquid of GeTe shows a high atomic mobility ($D = 10^6$ cm^2/s) down to temperatures very close to the glass transition temperature. Our calculated values of the viscosity as a function of temperature are consistent with a high fragility (fragility index∼100) of the supercooled liquid. The comparison between the calculated self-diffusion coefficient and the viscosity demonstrates a breakdown of the Stokes-Einstein relation below a crossover temperature of 700 K. These results support the experimental evidence of a breakdown of SER in the similar compound GST inferred by Orava et al. from ultrafast DSC measurements [53]. We repeated the same analysis for models of the amorphous phase overheated above T_g. Because of the fast heating rate in our simulations and in PCM as well, hysteretic effects are found, the overheated amorphous keeping a lower mobility than the supercooled liquid also above T_g. Still, $D \sim 5 \cdot 10^{-7}$ cm^2/s also in the overheated amorphous phase at 500 K. This feature is one of the keys to understand the origin of the high crystallization rate in phase change memories, in which the amorphous phase is heated above Tg. Actually the fast crystallization of GeTe and GeSbTe alloys has been ascribed to the similarity of the bonding topology in the amorphous and crystalline phases, as the must abundant rings in amorphous GeSbTe and GeTe are the four-membered ABAB rings (A=Ge/Sb and

B=Te) which are also the building blocks of the cubic crystalline phase. The presence of nanocavities in the amorphous phase has also been proposed as a structural feature that aids the alignment of the four-membered rings during crystallization. Our results suggest that in addition to the presence of these structural features in the amorphous phase, the crystallization is promoted also by a high atomic mobility just above T_g where, due to high supercooling, a large driving force ($\Delta\mu$) boosts the crystallization speed. A large atomic diffusivity coexists with a large viscosity due to the breakdown of SER. Therefore the self-diffusion coefficient to be used in the modeling of the crystallization process by classical nucleation theory can not be inferred from the expected viscosity and the use of SER, nor the crystallization speed can be extrapolated from measurements below T_g often reported in literature. Indeed the crystallization the amorphous phase at low temperatures below T_g of interest for data retention might take place in a different manner with respect to the crystallization of the of highly mobile overheated amorphous phase and of the supercooled liquid above T_g . Large scale simulations based on the NN potential hold the promise to shed light on the crystallization at low temperatures below Tg as well in the near future.

Finally, we studied the homogeneous crystallization of both supercooled liquid and overheated amorphous GeTe by means of direct molecular dynamics simulations. We observed crystallization on the timescale of 1 ns in models 4096-atom large in the temperature range 550-700 K. We observed an increase of the critical nucleus size with temperature as expected by classical nucleation theory with sizes that vary from 15 atoms at 500 K to 80 atoms at 700 K. We have estimated a speed of crystal growth in the range 0.5-4 m/s for temperatures in the range 550-700 K, comparable with recent experimental results on the similar GST compound by Orava *et al.* [53]. They observed speed of crystal growth is consistent with classical nucleation theory once our theoretical values of the self-diffusion coefficient and a difference in free energy between the liquid and the crystal are used. These results suggest that the high atomic mobility at high supercooling is indeed the origin of the fast crystallization in phase change materials.

Appendixes

6.2 Structural and vibrational properties of a-GeTe: fluctuations in small models

To assess the size of the fluctuations in the structural properties in a small cell, we here report the properties of a-GeTe for ten independent 216-atom models generated by quenching from the melt in DFT and NN simulations. The partial and total pair correlation functions, angle distribution functions and distribution of the coordination numbers are reported in Figs. 6.14,6.16,6.18 and 6.13,6.15,6.17 for NN and DFT models, respectively. Comparison with previous DFT data from Ref. [26] is also shown in Fig. 6.13.

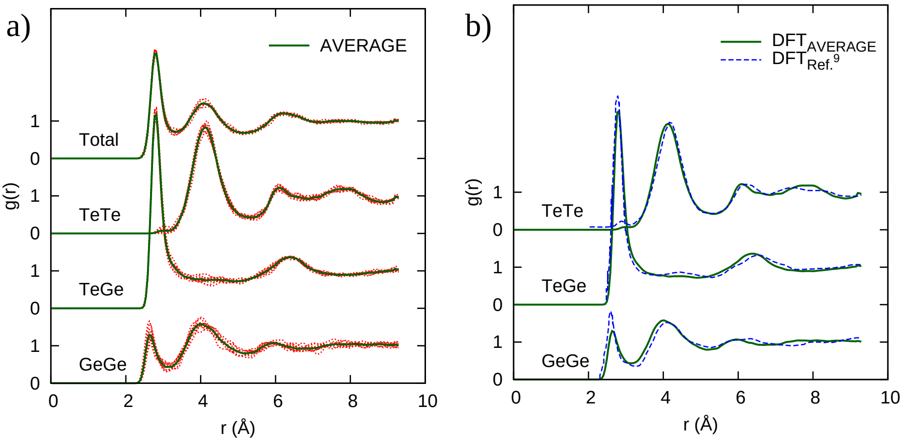

Figure 6.13: (a) Total and partial pair correlation functions of amorphous GeTe from ab initio molecular dynamics simulations at 300 K of ten independent 216-atom models. (b) Partial pair correlation functions average over the ten 216-atom models compared with previous DFT results with the same cell size and the same PBE functional from Ref. [26].

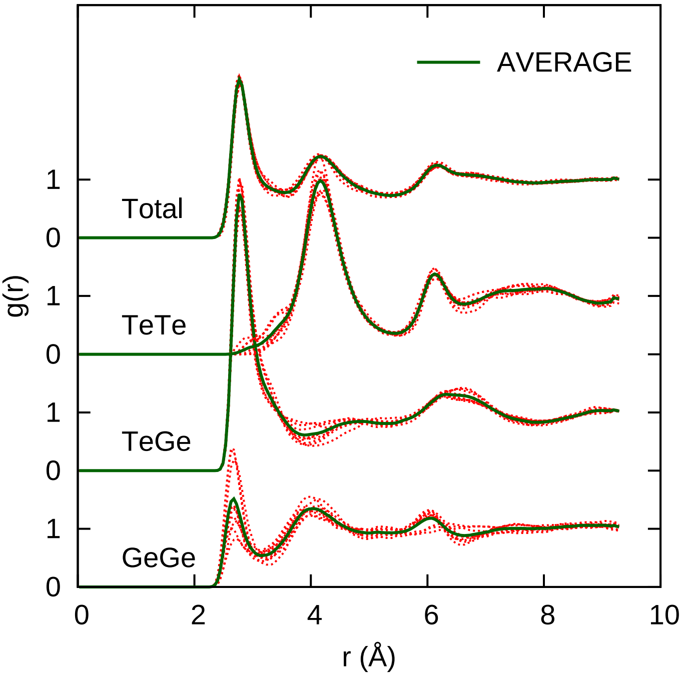

Figure 6.14: Total and partial pair correlation functions of amorphous GeTe from NN molecular dynamics simulations at 300 K of ten independent 216-atom models. Average values are also reported.

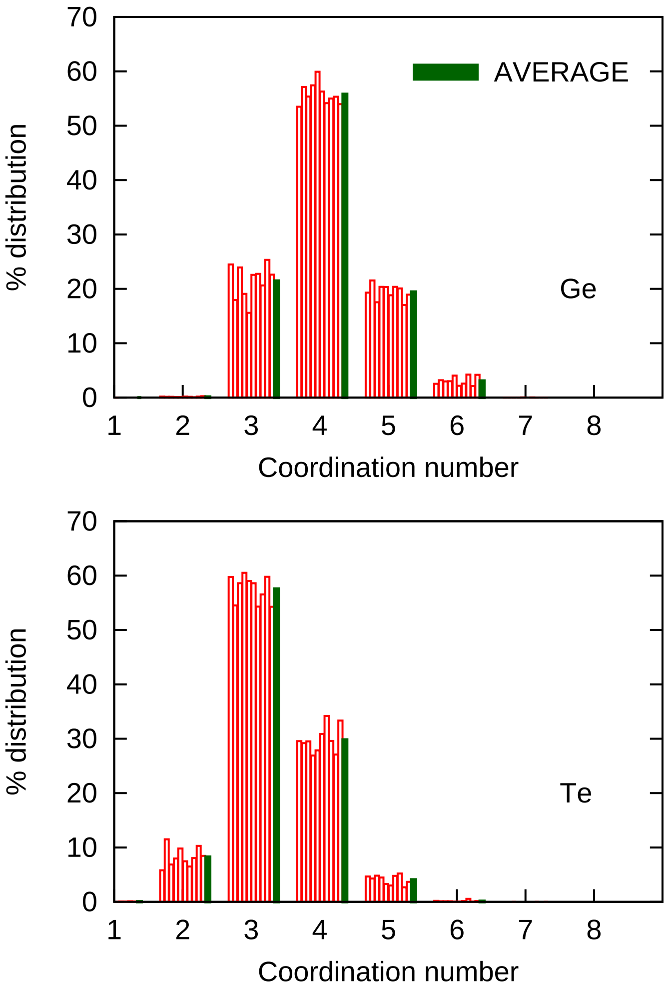

Figure 6.15: Distribution of coordination numbers of Ge and Te atoms of amorphous GeTe from DFT molecular dynamics simulations at 300 K of ten independent 216-atom models. Average values are also reported.

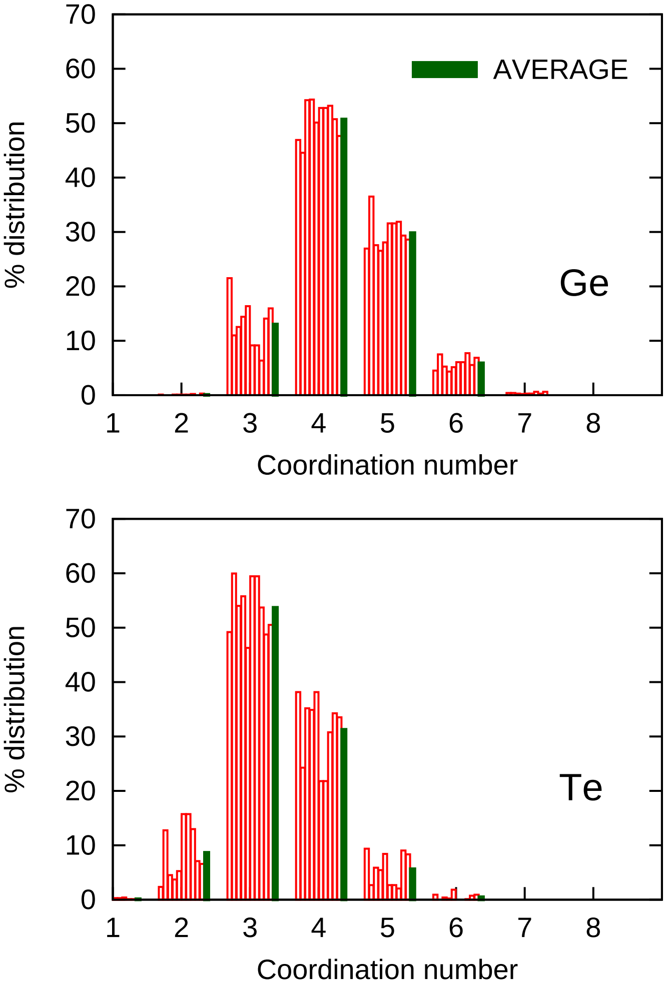

Figure 6.16: Distribution of coordination numbers of Ge and Te atoms of amorphous GeTe from NN molecular dynamics simulations at 300 K of ten independent 216-atom models. Average values are also reported.

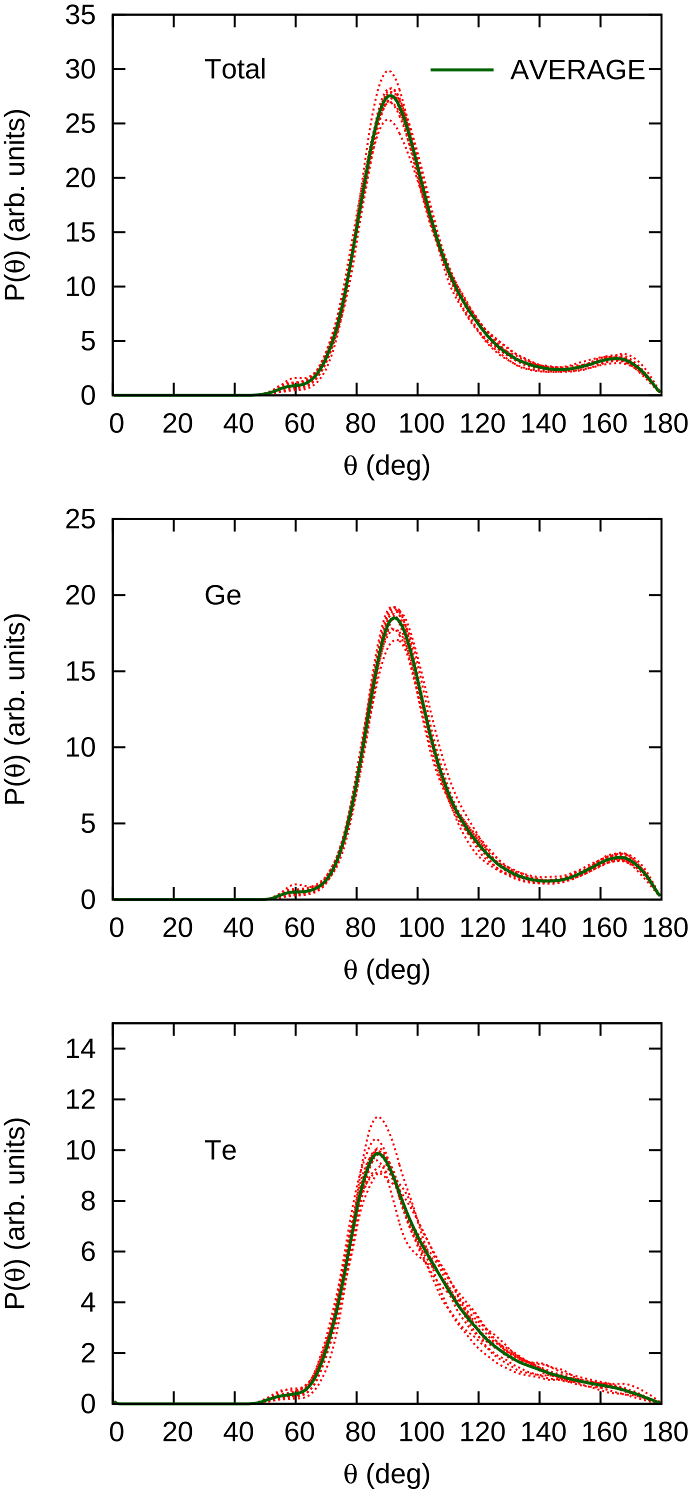

Figure 6.17: Total and partial angle distribution functions of amorphous GeTe from DFT molecular dynamics simulations at 300 K of ten independent 216-atom models. Average values are also reported. Partial distributions refer to X-Ge-Y and X-Te-Y

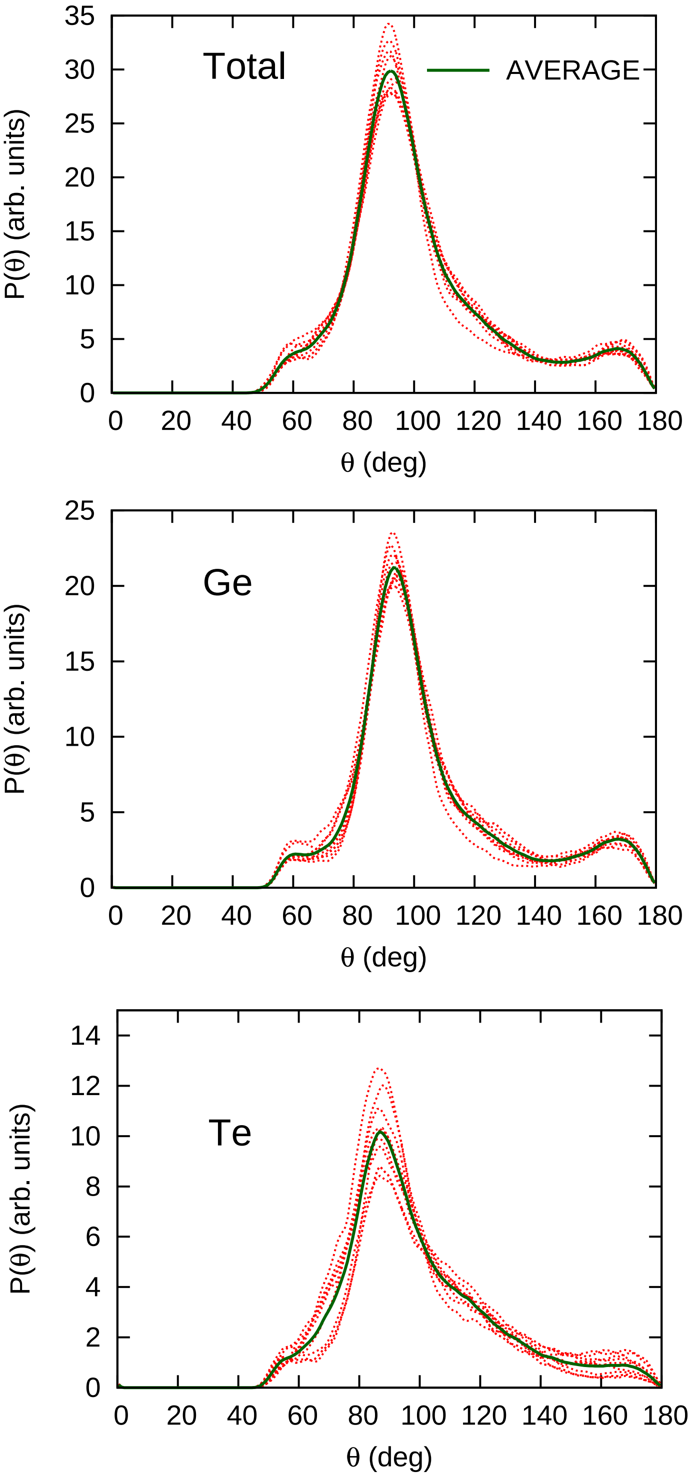

Figure 6.18: Total and partial angle distribution functions of amorphous GeTe from NN molecular dynamics simulations at 300 K of ten independent 216-atom models. Average values are also reported. Partial distributions refer to X-Ge-Y and X-Te-Y triplets (X,Y= Ge or Te).

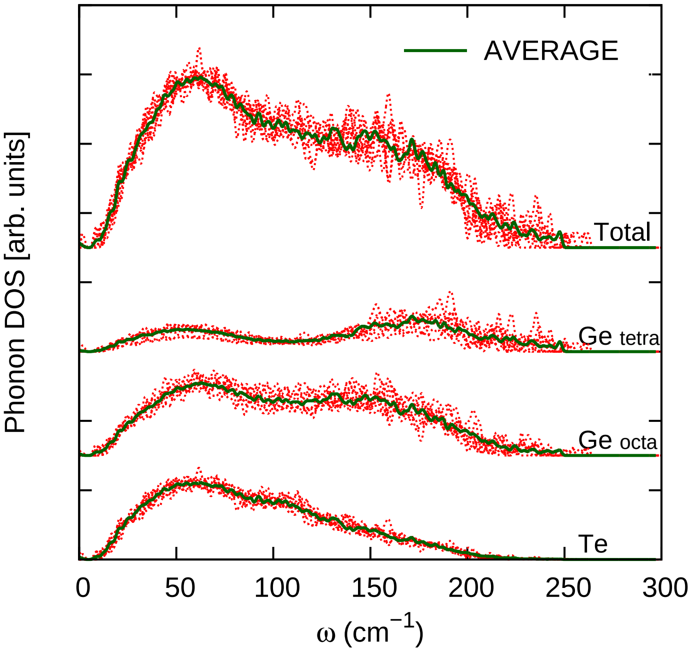

Figure 6.19: Phonon densities of states and projections on the different atomic species of ten indipendent 216-atom NN models. The averaged DOS is also shown.

6.3 Kolmogorov's theorem

Kolmogorov's theorem [160] states that any continuous real-valued function $f(x_1, x_2, ..., x_n)$ defined on $[0, 1]^n$, $n \geq 2$, can be represented in the form

$$f(x_1, x_2, ..., x_n) = \sum_{j=1}^{2n+1} g_i(y_i) \tag{6.7}$$

with

$$y_i = \sum_{j=1}^{n} \phi_{ij}(x_i) \tag{6.8}$$

where g_i are properly chosen continuous functions of one variable, and the ϕ_{ij} functions are continuos monotonically increasing functions independent of f. In other words, every continuos function of several variables (for a closed and bounded input domain) can be written as the superposition of a certain number of functions of one variable. It is has been shown [161] that Kolmogorov's theorem has no practical relevance in actual NN because g_j are usually far from being smooth, and even if one could be able to write g_j as the sum of infinte smooth functions, the finite number of adjustable parameters in Eq.6.7 would correspond to a finite number of degrees of freedom, while a generic continuos function has effectively infinitely many degrees of freedom. However, the debate about Kolmogorov's theorem is still ongoing and in any case this remarkable theorem casts the basis for a computational machinery that in principle -given an infinte numbers of parameters- is able to approximate any real valued continuos function: the feed forward neural network, which is the very core of many NN methods included the one we have chosen in order to construct our NN interatomic potential for GeTe.

Publications

- G. C. Sosso, S. Caravati, R. Mazzarello and M. Bernasconi, *“Raman spectra of cubic and amorphous* $Ge_2Sb_2Te_5$ *from first principles“*, Phys. Rev. B **83**, 134201 (2011)
- G. C. Sosso, G. Miceli, S. Caravati, J. Behler and M. Bernasconi, *“Neural network interatomic potential for the phase change material GeTe“*, Phys. Rev. B **85**, 174103 (2012)
- G. C. Sosso, J. Behler and M. Bernasconi, *“Breakdown of StokesEinstein relation in the supercooled liquid state of phase change materials“*, Phys. Stat. Sol. B **249**, 1880 (2012)
- G. C. Sosso, D. Donadio, S. Caravati, J. Behler and M. Bernasconi, *“Thermal transport in phase-change materials from atomistic simulations“*, Phys. Rev. B **86**, 104301 (2012)
- G. C. Sosso, J. Behler and M. Bernasconi, *“Breakdown of Stokes-Einstein relation in the phase change compound GeTe“*, (submitted)
- G. C. Sosso, G. Miceli, S. Caravati, J. Behler and M. Bernasconi, *“Heterogeneous and homogeneous crystallization of the phase change material GeTe via direct molecular dynamics simulations“*, (in preparation)
- S. Caravati, G. C. Sosso, M. Bernasconi and M. Parrinello, *“Phase transformation of hexagonal and cubic GST at high pressure“*, (in preparation)

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Marco Bernasconi, who tirelessly taught me everything I need and more, constantly pointing me in the right direction, wisely and uprightly leading me trough these three years with superhuman patience and unfaltering hope.

The very core of this work has been constructed thanks to the priceless help of Dr. Joerg Behler, who introduced me to the black magic of Neural Networks and kindly hosted me in his research group in Bochum for almost six months. All the best to him and to the whole RUB crew.

Yet another gracious German host, Dr. Davide Donadio enlightened me about thermal transport in amorphous systems in a very short time with swift accuracy and efficiency. Two months in his research group in Mainz were enough to learn a lot and have a lot fun as well. My grateful thanks to him, his beautiful family and to the whole MPIP crew.

A special thank is due to Dr. Sebastiano Caravati for helping us all, in particular with his mighty and apparently boundless computer skills. It has been something like six years since he started patiently teaching us the real Kung-Fu, and he will always be the greatest master.

Likewise, I'm deeply beholden to all the people who joined the U5 research group here in Milan, and particularly to Dr. Giacomo Miceli, Davide Mandelli, Davide Colleoni and Silvia Gabardi. Something more than just colleagues, they shared with me science as well as non conventional evenings, nights and (possibly) mornings. The same applies for ”the neighbors”: Francesca Boioli, Roberto Bergamaschin and Marco Salvalaglio, always available from discussions about the most intriguing details of my work and my favourite beer brands.

This should be no place to thank those few people I really care about. However, I have to mention my family, an impossible mixture of people who make me so proud, and I simply must include in these acknowledgments the three musketeers: Daniele Colombo, Stefano Bergantin and Umberto Martinez. They are roots, trees, rocks, falls, they are mountains and even more.

Last lines are for my little Irish fairy, who together with me survived to the tricky task of writing a PhD thesis while overcrowded by the incoming wedding. It's like they say: faith and trust, all you need is pixie dust.

This is the thesis of Via Cozzi 50, of its bloody trains, beetles, rats and strangers. This is the byproduct of many people who crossed my path and left their mark, carving a frame that undoubtedly made this thesis LEGEN -wait for it- DARY.

Bibliography

[1] M. Wuttig and N. Yamada, *Nat. Mat.* **6**, 824 (2007).

[2] A. Pirovano, A. L. Lacaita, A. Benvenuti, F. Pellizzer, and R. Bez, *IEEE Trans. Electron Devices* **51**, 452 (2007).

[3] A. L. Lacaita and D. J. Wouters, *Phys. Status Solidi A* **205**, 2281 (2008).

[4] S. Ovshinsky, *Phys. Rev. Lett.* **21**, 1450 (1968).

[5] D. Lencer, M. Salinga, and M. Wuttig, *Adv. Mater.* **23**, 2030 (2011).

[6] S. Raoux, W. Welnic, and D. Ielmini, *Chem. Rev.* **110**, 240 (2010).

[7] S. Caravati, M. Bernasconi, T. Kuhne, M. Krack, and M. Parrinello, *Appl. Phys. Lett.* **91**, 171906 (2007).

[8] J. Akola and R. Jones, *Phys. Rev. B* **76**, 235201 (2007).

[9] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).

[10] R. Mazzarello, S. Caravati, S. Angioletti-Uberti, M. Bernasconi, and M. Parrinello, *Phys. Rev. Lett.* **104**, 085503 (2010).

[11] W. Risk, C. Rettner, and S. Raoux, *Appl. Phys. Lett.* **94**, 101906 (2009).

[12] S. Zhang, J. He, T. Zhu, X. Zhao, and T. Tritt, *J. Non Cryst. Sol.* **355**, 79 (2009).

[13] S.-H. Lee, Y. Jung, and R. Agarwal, *Nature Nanotechnology* **2**, 626 (2007).

[14] M. J. Polking *et al.*, *Nanoletters* **11**, 1147 (2011).

[15] Y. He, D. Donadio, and G. Galli, *Appl. Phys. Lett.* **98**, 144101 (2011).

[16] C. Angell, *Science* **267**, 1924 (1995).

[17] J. Hegedus and S. Elliott, *Nature Mater.* **7**, 399 (2008).

[18] R. Bez, *Mat. Sci. Sem. Proc.* **7**, 349 (2004).

[19] M. Lankhorst, B. Ketelaars, and R. Wouters, *Nat. Mat.* **4**, 347 (2005).

[20] T. Morikawa *et al.*, *Electron Devices Meeting* **1**, 307 (2007).

[21] W. Klemm and G. Frischmuth, *Z. Anorg. Chem.* **218**, 249 (1934).

[22] D. I. Bletskan, *Journal of Ovonic Research* **1**, 53 (2005).

[23] J. Enderby, *Amorphous and Liquid Semiconductors* (Plenum, New York, 1974).

[24] J. Raty, J. Gaspard, M. Bionducci, R. Ceolin, and R. Bellissent, *J. Non-Cryts. Solids* **250-252**, 277 (1999).

[25] J. Raty, V. Godlevsky, J. Gaspard, C. Bichara, M. Bionducci, R. Bellissent, R. Ceolin, J. Chelikowsky, and P. Ghosez, *Phys. Rev. B* **65**, 115205 (2002).

[26] J. Akola and R. Jones, *Phys. Rev. Lett.* **100**, 205502 (2008).

[27] S. Kohara *et al.*, *Appl. Phys. Lett.* **89**, 201910 (2006).

[28] J. Raty, V. Godlevsky, P. Ghosez, C. Bichara, J. Gaspard, and J. Chelikowsky, *Phys. Rev. Lett.* **85**, 1950 (2000).

[29] A. Kolobov and P. F. and. J. Tominaga, *Appl. Phys. Lett.* **82**, 382 (2003).

[30] J. Errington and P. Debenedetti, *Nature* **409**, 318 (2001).

[31] L. Giacomazzi, P. Umari, and A. Pasquarello, *Phys. Rev. B* **74**, 155208 (2006).

[32] D. Franzblau, *Phys. Rev. B* **44**, 4925 (1991).

[33] S. Caravati, M. Bernasconi, and M. Parrinello, *Phys. Rev. B* **81**, 014201 (2010).

[34] J. Akola, J. Larrucea, and R. Jones, *Phys. Rev. B* **83**, 094113 (2011).

[35] J. Kalikka, J. Akola, R. Jones, S. Kohara, and T. Usuki, *E/PCOS Proceedings* **1**, 1 (2011).

[36] K. Andrikopoulos *et al.*, *J. Phys. Chem. Sol.* **68**, 1074 (2007).

[37] S. Caravati, M. Bernasconi, T. Khune, M. Krack, and M. Parrinello, *Phys. Rev. Lett.* **102**, 205502 (2009).

[38] W. Phillips, *J. Low Temp. Phys.* **7**, 351 (1972).

[39] N. Mott, *Philos. Mag.* **13**, 93 (1966).

[40] P. Allen and J. Feldman, *Phis. Rev. B* **48**, 12581 (1993).

[41] P. Anderson, *Phys. Rev.* **109**, 1942 (1958).

[42] P. Allen, J. Feldman, J. Fabian, and F. Wooten, *Philosop. Magaz. B* **79**, 1715 (1999).

[43] P. D. Benedetti, *Metastable Liquids: Concepts and Principles* (Princeton University Press, Princeton, 1996).

[44] A. Onuki, *Phase Transition Dynamics* (Cambridge University Press, Cambridge, 2002).

[45] S. Prestipino, A. Laio, and E. Tosatti, *Phys. Rev. Lett.* **108**, 225701 (2012).

[46] D. Moroni, P. R. ten Wolde, and P. Bolhuis, *Phys. Rev. Lett.* **94**, 235703 (2005).

[47] S. Raoux and M. Wuttig, *Phase Change Materials: Science and Aplications* (Springer, New York, 2009).

[48] M. Ediger, M. Harrowell, and P. Yu, *J. Chem. Phys.* **128**, 034709 (2008).

[49] C. Thompson and F. Spaepen, *Acta Metall.* **27**, 1855 (1979).

[50] L. Battezzati and A. Greer, *J. Mater. Res.* **3**, 570 (1988).

[51] V.weidenhof, I. Friedrich, S. Ziegler, and M. Wuttig, *J. Appl. Phys* **89**, 3168 (2001).

[52] J. D. H. B.J. Kooi, *J. Appl. Phys* **95**, 4714 (2004).

[53] J. Orava, A. Greer, B. Gholipour, D. Hewak, and C. Smith, *Nature Mater.* **11**, 279 (2012).

[54] W. McCulloch and W. Pitts, *Bull. Math. Biophys.* **5**, 115 (1943).

[55] R. Hecht-Nielsen, *Proc. Int. Joint Conf. Neural Networks* **1**, 593 (1989).

[56] A. Waibel, *Neural Comp.* **1**, 39 (2008).

[57] A. Rajavelu, M. Musavi, and M. Shirvaikar, *Neural Networks* **2**, 387 (1989).

[58] C. Peterson and B. Söderberg, *Int. J. Neural Systems* **1**, 3 (1989).

[59] A. Chen and M. Leung, *J. Forecasting* **24**, 403 (2005).

[60] C. Marzban and A. Witt, *Weather and Forecasting* **16**, 600 (2001).

[61] K. Funahashi, *Neural Networks* **2**, 183 (1989).

[62] G. Cybenko, *Math. Contr. Sign. Sys.* **2**, 303 (1989).

[63] W. Press, S. Teukolsky, W. Vetterling, and P. Flannert, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1992).

[64] J. Ischtwan and M. Collins, *J. Chem. Phys.* **100**, 8080 (1994).

[65] A. Bartok, M. Payne, R. Kondor, and G.Csanyi, *Phys. Rev. Lett.* **104**, 136403 (2010).

[66] R. Bellman, *Adaptive Control Processes: A Guided Tour* (Princeton University Press, New Jersey, 1961).

[67] H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, *J. Phys. Chem. A* **102**, 4596 (1998).

[68] K. Cho, K. No, and H. Scheraga, *J. Mol. Struct.* **641**, 77 (2002).

[69] H. Le and L. Raff, *J. Phys. Chem. A* **114**, 45 (2010).

[70] S. Lorenz, A. Groß, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).

[71] J. Ludwig and D. Vlachos, *J. Chem. Phys.* **127**, 154716 (2007).

[72] J. Behler, K. Reuter, and M. Scheffler, *Phys. Rev. B* **77**, 115421 (2008).

[73] S. Hobday, R. Smith, and J. Belbruno, *Model. Simul. Mater. Sci. Eng.* **7**, 397 (1999).

[74] J. Tersoff, *Phys. Rev. Lett.* **56**, 632 (1986).

[75] A. Bholoa, S. Kenny, and R. Smith, *Nucl. Instrum. Methods Phys. Res. Sect. B* **255**, 1 (2006).

[76] E. Sanville, A. Bholoa, R. Smith, and S. Kenny, *J. Phys. Cond. Matt.* **20**, 285219 (2008).

[77] N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* **83**, 153101 (2011).

[78] M. Allen and D. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1989).

[79] A. Gelb, *Applied Optimal Estimation* (MIT Press, Cambridge, MA, USA, Cambridge, 1974).

[80] T. Blank and S. Brown, *J. Chemometrics* **8**, 391 (1994).

[81] J. Witkoskie and D. Doren, *J. Chem. Theory Comput.* **1**, 14 (2005).

[82] D. Frenkel and B. Smit, *Understanding Molecular Simulations* (Academic Press, San Diego, 1996).

[83] W. Swope, H. Andersen, P. Berens, and K. Wilson, *J. Chem. Phys.* **76**, 637 (1982).

[84] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, New York, 2000).

[85] D. Marx and J. Hutter, *Modern Methods and Algorithms of Quantum Chemistry* (J. von Neumann Institute for Computing, NIC series, 2000).

[86] H. Hellmann, *Z. Physics* **85**, 180 (1933).

[87] R. Feynman, *Phys. Rev.* **56**, 340 (1939).

[88] S. Kirkpatrick, C. Gelatt, and M. Vecchi, *Science* **220**, 671 (1983).

[89] R. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).

[90] R. Dreizler and E. Gross, *Density Functional Theory* (Springer-Verlag, New York, 1990).

[91] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, 864 (1964).

[92] W. Kohn and L. Sham, *Phys. Rev.* **140**, 1133 (1965).

[93] M. Payne, *Rev. Mod. Phys.* **64**, 1045 (1992).

[94] D. Ceperley and B. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).

[95] J. Perdew, K. Burke, and M. Erzenhof, *Phys. Rev. Lett.* **77**, 3685 (1996).

[96] S. Kurth, J. Perdew, and P. Blaha, *Int. J. Quantum. Chem.* **75**, 889 (1999).

[97] J. Perdew, A. Ruzsinszky, J. Tao, V. Staroverov, G. Scuseria, and G. Csonka, *J. Chem. Phys.* **123**, 62201 (2005).

[98] J. Perdew, *Electronic Structure of Solids* (Akademie Verlag, Berlin, 1991).

[99] J. Akola and R. Jones, *Phys. Status Solidi B* **249**, 1851 (2012).

[100] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, *Phys. Rev. Lett.* **100**, 136406 (2008).

[101] J. Tao, J. Perdew, V. Staroverov, and G. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).

[102] G. Lippert, J. Hutter, and M. Parrinello, *Mol. Phys.* **92**, 477 (1997).

[103] J. Vandevondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Comp. Phys. Comm.* **167**, 103 (2005).

[104] M. Frigo and S. Johnson, *Proceedings of the IEEE* **93**, 216 (2005).

[105] S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).

[106] P. Pulay, *J. Comput. Chem.* **3**, 556 (1982).

[107] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).

[108] T. Kune, M. Krack, F. Mohamed, and M. Parrinello, *Phys. Rev. Lett.* **98**, 066401 (2007).

[109] J. Kolafa, *J. Comput. Chem.* **25**, 335 (2004).

[110] M. Krack and M. Parrinello, in High Performance Computing in Chemistry J. Grotendorst (ed.), NIC vol. 25, 29-51 (2004); http://cp2k.berlios.de.

[111] D. Kofke, *J. Chem. Phys.* **98**, 4149 (1993).

[112] J. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).

[113] J. Anwar, D. Frenkel, and M. Noro, *J. Chem. Phys.* **118**, 728 (2003).

[114] J. Johnson, J. Zollweg, and K. Gubbins, *Mol. Phys.* **78**, 591 (1993).

[115] S. Grimme, *J. Comput. Chem.* **27**, 1787 (2006).

[116] V. Glazov and O. Shelicov, *Sov. Phys. Semicond.* **18**, 411 (1984).

[117] M. Alinchenko, A. Anikeenko, N. Medvedev, V. Voloshin, M. Mezei, and P. Jedlovszky, *J. Phys. Chem. B* **108**, 19056 (2004).

[118] N. Medvedev, V. Voloshin, V. Luchnikov, and M. Gavrilova, *J. Comp. Chem.* **27**, 1676 (2006).

[119] R. Martoňák, A. Laio, and M. Parrinello, *Phys. Rev. Lett.* **90**, 075503 (2003).

[120] E. Spreafico, S. Caravati, and M. Bernasconi, *Phys. Rev. B* **83**, 144205 (2011).

[121] CPMD, Version 3.11, IBM Corp., 1990-2011, MPI fur Festkorperforschung, Stuttgart 1997-2001, http://www.cpmd.org.

[122] H. J. Monkhorst and J. Pack, *Phys. Rev. B* **13**, 5188 (1976).

[123] P. G. et al., *J. Phys. Cond. Matt.* **21**, 395502 (2009).

[124] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).

[125] RuNNer: A Neural Network Code for High-Dimensional Potential-Energy Surfaces, Jorg Behler, Lehrstuhl fur Theoretische Chemie, Ruhr-Universitat Bochum, Germany.

[126] W. Smith and T. Forester, *J. Mol. Graph.* **14**, 136 (1996).

[127] H. Berendsen, J. Postma, W. van Gunsteren, A. DiNola, and J. Haak, *J. Chem. Phys.* **81**, 3684 (1984).

[128] J. Goldak, C. Barrett, D. Innes, and W. Youdelis, *J. Chem. Phys.* **44**, 3323 (1966).

[129] G. Ghezzi, J. Raty, S. Maitrejean, A. Roule, E. Elkaim, and F. Hippert, *Appl. Phys. Lett.* **99**, 151906 (2011).

[130] F. Murnaghan, *PNAS* **30**, 244 (1944).

[131] D. Lencer, M. Salinga, B. Grabowski, T. Hickel, J. Neugebauer, and M. Wuttig, *Nat. Mat.* **7**, 972 (2008).

[132] R. Shaltaf, E. Durgun, J. Raty, P. Ghosez, and X. Gonze, *Phys. Rev. B* **78**, 205203 (2008).

[133] G. Sosso, S. Caravati, R.Mazzarello, and M. Bernasconi, *Phys. Rev. B* **83**, 134201 (2011).

[134] G. Sosso, G. Miceli, S. Caravati, J. Behler, and M. Bernasconi, *Phys. Rev. B* **85**, 174103 (2012).

[135] W. Hoover, *Computational Statistical Mechanics* (Elsevier, Amsterdam, 1991).

[136] Y. Lee, R. Biswas, C. Soujoulis, C. Wang, C. Chan, and K. Ho, *Phys. Rev. B* **43**, 6573 (1991).

[137] E. Helfand, *Phys. Rev.* **119**, 1 (1960).

[138] J. Che, T. Cagin, W. Deng, and W. Goddard, *J. Chem. Phys.* **113**, 6888 (2000).

[139] G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).

[140] P. Nath and K. Chopra, *Phys. Rev. B* **10**, 3412 (1974).

[141] G. Srivastava, *The Physics of Phonons* (Adam Hilger, Bristol, 1990).

[142] N. Ashcroft and N. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).

[143] R. Peierls, *Quantum Theory of Solids* (Clarendon, Oxford, 1956).

[144] D. Cahill, S. Watson, and R. Pohl, *Phys. Rev. B* **46**, 6131 (1992).

[145] G. Sosso, D. Donadio, S. Caravati, J. Behler, and M. Bernasconi, *Phys. Rev. B* **86**, 104301 (2012).

[146] J. Kalb, F. Spaepen, T. L. Pedersen, and M. Wuttig, *J. App. Phys.* **94**, 4908 (2003).

[147] M. Micolaut, R. Vuilleumier, and C. Massobrio, *Phys. Rev. B* **79**, 214205 (2009).

[148] D. Derewnicka, P. Zielinski, and H. Davies, *J. Mat. Sci. Lett.* **1**, 87 (1982).

[149] S. Raoux, B. M. H. Cheng, and J. Jordan-Sweet, *Appl. Phys. Lett.* **95**, 143118 (2009).

[150] J. Mauroa, Y. Yueb, A. Ellisona, P. Guptac, and D. Allan, *Proc. Natl. Acad. Sci.* **106**, 19780 (2009).

[151] M. Cohen and G. Grest, *Phys. Rev. B* **20**, 1077 (1979).

[152] R. Bohmer, K. Ngai, C. Angell, and D. J. Plazek, *J. Chem. Phys.* **99**, 4201 (1993).

[153] B. Dunweg and K. Kremer, *J. Chem. Phys.* **99**, 6983 (1993).

[154] R. Zallen, *The Physics of Amorphous Solids* (Wiley-VCH, Weinheim, 1993).

[155] G. Sosso, J. Behler, and M. Bernasconi, *Phys. Stat. Sol. B* **249**, 1880 (2012).

[156] P. Steinhardt, D. Nelson, and M. Ronchetti, *Phys. Rev. B* **28**, 784 (1983).

[157] F. Auer and D. Frenkel, *Adv. Polym. Sci.* **173**, 149 (2005).

[158] J. Rudnick and G. Gaspari, *J. Phys A. Math. Gen.* **19**, L191 (1986).

[159] C. Rycroft, *Chaos* **19**, 041111 (2009).

[160] A. Kolmogorov, *Doklady Akademiia Nauk SSSR* **114**, 953 (1957).

[161] F. Girosi and T. Poggio, *Neural Computation* **1**, 465 (1989).