Studies of phase-change memory materials using machine-learned interatomic potentials



Yuxing Zhou

Exeter College

A thesis submitted to the University of Oxford for the degree of Doctor of Philosophy

Inorganic Chemistry Laboratory

University of Oxford

Trinity 2024

Abstract

In the era of big data, the growing demand for advanced data storage and processing technologies has highlighted the use of phase-change materials (PCMs), particularly Ge–Sb–Te (GST) ternary compounds, as candidates for next-generation data storage and neuromorphic computing chips. Quantum-accurate simulations (e.g., ab initio molecular dynamics simulations) have played a crucial role in understanding the properties of PCMs, but their expensive computational costs have limited their application to the length and time scales of real PCM-based devices, hindering atomic-scale understanding of switching mechanisms in the devices.

To overcome these limitations, this thesis explores the use of machine-learned (ML) interatomic potentials to study PCMs at length and time scales relevant to real-world devices. Two PCM systems are investigated: elemental tellurium (Te), a phase-change selector material, and GST alloys, the flagship phase-change memory materials with greater structural and chemical complexity. This thesis proposes a domain-knowledge-based iterative training protocol to construct the training databases, and employs two fitting frameworks—Gaussian approximation potential (GAP) and atomic cluster expansion (ACE)—to develop ML potentials with optimised hyperparameters.

The thesis first studies the microscopic mechanisms of crystallisation and chirality transfer in Te. Bespoke GAP and ACE potentials were developed and validated against quantum-accurate simulations. Large-scale ACE- and GAP-driven molecular dynamics simulations reveal that Te initially forms disordered cubic-like seeds during nucleation, which transform into chiral helices. Frequent chirality transfer and significant grain-boundary ordering were observed at elevated temperatures, mediated by forming cubic-like motifs. Large-scale simulations, comparable to real-world device dimensions, uncovered the full-loop switching behaviours of Te-based devices and demonstrated the impact of grain morphology on switch-on voltages.

For the structurally and chemically more complex GST alloys, a diverse training database was constructed, leading to a quantum mechanically accurate GAP potential for all compositions in the pseudo binary tie-line of Sb_2Te_3 and GeTe. Based on the extension of the reference database and optimised hyperparameters, a faster and equally accurate ACE potential was obtained. These ML potentials were used to simulate realistic programming operations in GST-based devices, e.g., iterative RESET, cumulative SET, and non-isothermal heating processes. Full-loop atomistic simulations on two device geometries—confined-type and mushroom-type—were carried out, demonstrating the nanosecond-scale phase transitions in GST-based devices. A temperature-dependent competition between nucleation and crystal growth was observed: lower crystallisation temperatures favour nucleation and polycrystalline states with many grains, while higher crystallisation temperatures promoted rapid crystal growth, resulting in fewer and larger grains.

In conclusion, this thesis provides atomic-level insights into the fundamental properties and switching mechanisms of structurally and chemically complex phase-change solids using ML-driven simulations, which can contribute to the further development of ML potentials and the study of other complex functional materials.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Volker L. Deringer. His mentorship over the past four years, especially those during the hardest Covid time, has been absolutely valuable to my DPhil life. His insightful guidance, unwavering support, and profound knowledge have shaped not only my research but also my development as a chemist. Without his constant encouragement and supports, I could not have overcome the countless challenges I faced during my DPhil journey. There are so many details that he has done for me, which I cannot write down on one single page. I am so proud of the achievements we have gained together in these years. I am deeply thankful for his patience, wisdom, and the trust he placed in me.

I am also immensely grateful to my collaborators, each of whom has played a significant role in my research over the past four years. I would like to thank Prof. Wei Zhang, a leading expert in phase-change materials. His insights into critical questions in the field were indispensable to many of my works. His advice and contributions have always been valuable to me. I also want to express my heartfelt thanks to Prof. Stephen R. Elliott, whose energy and creativity have consistently sparked new ideas and inspired me to push boundaries. His unwavering patience and great enthusiasm made working together a true pleasure. Prof. En Ma has also had a tremendous impact on my research. His distinguished expertise in disordered systems and his visionary ideas were a continuous source of inspiration.

Special thanks go to Daniel F. Thomas du Toit, who has been a key figure in our joint efforts. His technical expertise was instrumental to the success of many of our projects, and I am deeply appreciative of the many papers we worked on together. I would also like to thank Will Kirkpatrick. It is a huge pity that we did not meet in person due to the Covid. His contribution to the phosphorus project during his time as a Part-II student was invaluable.

To all of my group members—Tom, Joe, Daniel, Zak, John, Louise, Zoe, Shixuan, Natascia, Litong, Bianca, Bodoo, Stephen, Yuanbin, Andy, Chiheb, and all others who have been part of the Deringer group—thank you for the stimulating discussions and great supports. I will always remember the fun time we spent together. Your presence has truly transformed my life in Oxford. Your friendship and encouragement have made my DPhil years an unforgettable experience.

Finally, I owe a special thanks to my family for their unconditional love and support, especially during the most challenging times of my studies. To my girlfriend, Jieling Tan, we have been through the hardest times, especially during the first two years of my DPhil. Despite the challenges we faced, we never gave up. All the video calls across two countries stand as a true testament to our heartfelt bond. Your unwavering encouragement and belief in me have been my anchor throughout this journey. I could not have achieved this without you.

Key publications

[1] **Y. Zhou**, S. R. Elliott, D. F. Thomas du Toit, W. Zhang*, V. L. Deringer*, The pathway to chirality in elemental tellurium, preprint at arXiv:2409.03860 (2024).

[2] D. F. Thomas du Toit, **Y. Zhou**, V. L. Deringer*, Hyperparameter optimization for atomic cluster expansion potentials, J. Chem. Theory Comput., 20, 10103–10113 (2024).

[3] **Y. Zhou**, W. Zhang*, E. Ma, V. L. Deringer*, Device-scale atomistic modelling of phase-change memory materials, Nat. Electron., 6, 746–754 (2023).

[4] **Y. Zhou**, S. R. Elliott, V. L. Deringer*, Structure and bonding in amorphous red phosphorus, Angew. Chem. Int. Ed., 62, e202216658 (2023).

[5] **Y. Zhou**, W. Kirkpatrick, V. L. Deringer*, Cluster fragments in amorphous phosphorus and their evolution under pressure, Adv. Mater., 34, 2107515 (2022).

[6] Y. Xu#, **Y. Zhou**#, X. D. Wang, W. Zhang*, E. Ma, V. L. Deringer, R. Mazzarello*, Unraveling Crystallization mechanisms and electronic structure of phase-change materials by large-scale ab initio simulations, Adv. Mater., 34, 2270084 (2022).

# Authors who contributed equally to the work.

* Corresponding authors.

Contents

1 Introduction **1**

1.1 Overview 1

1.2 Phase-change materials and devices 2

1.3 Working mechanisms of phase-change materials 18

1.4 Ab initio atomistic simulations of phase-change materials 27

1.5 ML-driven atomistic simulations of phase-change materials 43

1.6 Challenges for atomic-scale modelling of PCM-based devices 50

1.7 Going beyond atomic-scale simulations 53

2 Thesis outline **58**

3 Methods **60**

3.1 Machine-learned interatomic potentials 60

3.2 Generation of training database 74

3.3 Hyperparameter optimisation using XPOT 82

3.4 The applications of ML potentials in the study of complex solids . . . 88

3.5 Quantification of per-atom crystallinity 107

4 Tellurium-based phase-change selector materials **119**

4.1 Chirality in elemental tellurium 119

4.2 The role of transient cubic-like motifs 121

4.3 Fitting ML potentials for tellurium 124

4.4 Crystallisation of Te at the atomic scale 131

4.5 Boundary ordering and chirality transfer . . . 142
4.6 Switching in Te selector devices . . . 147
4.7 Conclusion . . . 153

5 ML-driven modelling of Ge–Sb–Te phase-change materials 155
5.1 Atomistic modelling of Ge–Sb–Te alloys . . . 155
5.2 Fitting ML potentials for Ge–Sb–Te PCMs . . . 157
5.3 Validation of ML potentials . . . 166
5.4 Modelling conditions to represent devices . . . 176
5.5 Towards a complete reference database . . . 188
5.6 Conclusion . . . 192

6 Unlocking full-loop device-scale simulations 194
6.1 Towards full-loop device simulations . . . 194
6.2 Unlocking billion-atom simulations . . . 195
6.3 Simulating full-loop operations of crossbar devices . . . 200
6.4 Simulating full-loop operations of mushroom-type devices . . . 205
6.5 Conclusion . . . 207

7 Conclusion and outlook 208
7.1 Conclusion . . . 208
7.2 Outlook . . . 210

Bibliography 211

Chapter 1

Introduction

1.1 Overview

The global demand for digital data storage has grown explosively in recent decades. With the rapid development of data-driven applications (e.g., autonomous vehicles, large-language models, and supercomputing), the need for high-performance storage and computing technologies has greatly increased. Phase-change materials (PCMs) have been regarded as one of the most promising candidates to cope with these challenges. In particular, PCMs can be used in universal storage and memory, as well as emerging neuromorphic computing chips [1].

PCMs have two thermally stable states under ambient conditions, the amorphous state and the crystalline state. When annealed at high temperatures (above the melting point), PCMs turn into disordered liquid. A subsequent fast quench freezes the disordered network and forms an amorphous structure. When heating the amorphous PCMs at elevated temperatures close to the glass transition temperature (≈600 K), they will crystallise and switch back to the crystalline state. In PCM-based devices, external pulses (e.g., electric or laser pulses) trigger the fast phase transition under different programming conditions (e.g., different pulse widths and amplitudes). Given the large structural and chemical complexity in PCMs, understanding the fundamental properties of PCMs and the realistic programming processes of PCM-based devices at atomic scale is of great importance to design new memory devices with superb performance.

In contrast to the abundant experimental studies on crystalline PCMs, atomistic simulations have been a useful theoretical tool to model the structural properties and switching processes of PCMs. In particular, ab initio molecular dynamics (AIMD) simulations have been widely used to investigate their different properties and to

reveal the microscopic mechanism of phase transition processes. However, AIMD simulations are exceedingly expensive, with a maximum simulation size of ≈1,000 atoms, which hinders the atomistic modelling of real-world PCM-based devices.

Due to the fast development of machine-learned (ML) interatomic potentials, large-scale atomistic modelling has become feasible. After careful validation, ML potentials can simulate structural models containing hundreds of thousands of atoms within only several days using state-of-the-art supercomputers. To enable atomistic modelling for PCM-based devices, good strategies are required to obtain accurate and efficient ML potentials, and reasonable simulation protocols are needed to model the complex programming conditions in PCM-based devices.

In this chapter, an overview of PCMs and PCM-based devices will first be given, including fundamental properties of PCMs and working mechanisms of PCM-based devices. This is followed by a comprehensive discussion of atomistic simulations of PCMs, in particular the widely used AIMD simulations and newly emerging ML-driven molecular dynamics (MD) simulations. After reviewing the state-of-the-art studies of PCMs using atomistic simulations, challenges for device-scale modelling, which closely mimic the working conditions and programming processes in real-world devices, will be discussed. Lastly, the outline of the remaining chapters in this thesis will be presented.

1.2 Phase-change materials and devices

In 1968, Ovshinsky reported his finding of a rapid and reversible transition, triggered by an external electric field, between a highly resistive and conductive states in a chalcogenide, $Ge_{10}Si_{12}As_{30}Te_{48}$ [2]. He proposed the potential use for data storage, which led to the initial research on PCMs. In the following decade, much attention was paid to understanding the electrical switching and nonvolatile memory capabilities of these phase-change chalcogenides. Note that being nonvolatile refers to a type of memory or storage, e.g., solid-state disks (SSD) and hard-disk drives (HDD), which can retain data even when the external power is turned off, in stark contrast

to conventional silicon-based volatile memory and storage devices, e.g., dynamic random access memory (DRAM) and static random access memory (SRAM).

In the late 20th century, researchers from IBM found the binary compound GeTe, a simple compound that could undergo rapid phase changes [3]. Yamada et al. discovered various ternary compounds, such as $Ge_2Sb_2Te_5$, $Ge_1Sb_2Te_4$, $Ge_8Sb_2Te_{11}$, located on the pseudo-binary line between GeTe and Sb_2Te_3 of the Ge–Sb–Te (GST) system (referred to as GST alloys; Fig. 1.1) [4]. Given the pronounced optical contrast, these ternary compounds have fast crystallisation speeds, good thermal stability, and excellent cycling endurance, serving as good candidates for PCMs-based rewritable optical discs. A few years later, Sb_2Te alloys and Sb alloys doped with other elements were considered the main candidate materials for rewritable optical media products [5].

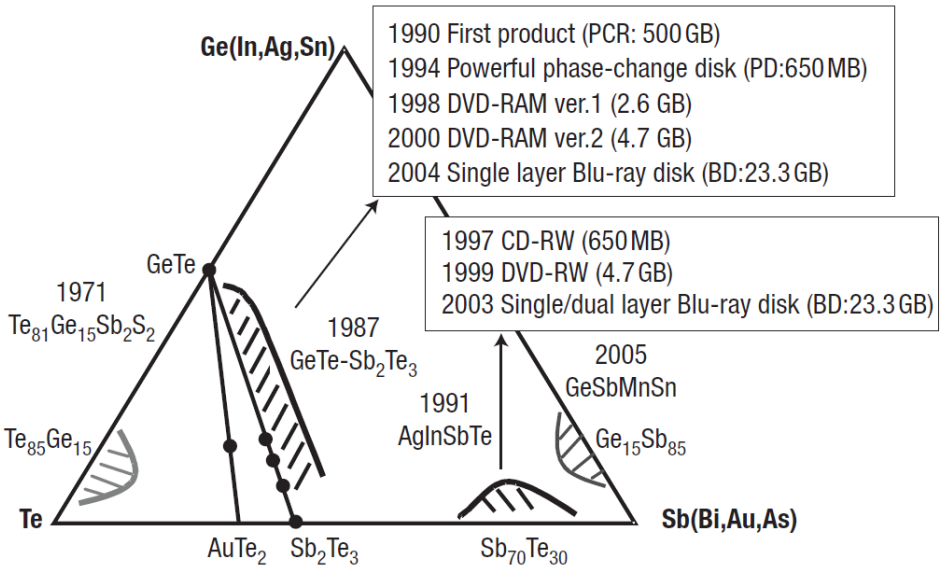


Figure 1.1: Ge–Sb–Te ternary diagram, depicting different phase-change alloys, the year of discovery and their use in different optical storage products. The figure is taken with permission from Ref. 6. Copyright 2006, Springer Nature Limited.

On the other hand, many companies and research institutions focused on GST alloys and doped Sb–Te alloys for their potential use as electronic devices. In 2005, Lankhorst et al. achieved a fast programming (i.e., writing and erasing operations) speed of less than 50 nanoseconds using doped Sb–Te alloys [5], which stimulated

the idea of using phase-change storage materials for universal memory [7]. In the following decade, GST alloys, especially the $Ge_2Sb_2Te_5$ alloy, were extensively studied and successfully used in the 3D XPoint devices.

As briefly summarised in Fig. 1.1, there are three main types of phase-change materials in the GST ternary diagram: alloys located on the GeTe-Sb_2Te_3 pseudo-binary line, the pristine and doped Sb_2Te compounds, and Sb and its doped systems. Among these, the alloys located on the GeTe-Sb_2Te_3 pseudo-binary line are the most studied. Due to better performance as compared to other PCMs, GST alloys have been widely used in phase-change storage devices (including both electronics and photonics). The Sb_2Te and Sb doped systems (such as $Ag_4In_3Sb_{67}Te_{26}$ and $Ge_{15}Sb_{85}$) are mainly used in rewritable optical disc applications [5]. Additionally, other unconventional phase-change materials like GaSb and In_3SbTe_2 are also used in phase-change storage electronic devices. This thesis will mostly focus on the study and discussion of GST alloys.

1.2.1 Structural properties of GST alloys

As the flagship PCM, GST alloys have been widely studied, both experimentally and theoretically, and implemented in high-performance memory applications. One notable example is the implementation of GST alloys in the 3D XPoint technology, released by Intel and Micron [8]. Among all GST compositions located in the pseudo binary line between Sb_2Te_3 and GeTe, $Ge_1Sb_2Te_4$ and $Ge_2Sb_2Te_5$ have attracted the most attention. Comprehensive understanding of the structural properties and the relationship between microscopic structure and macroscopic performance of GST alloys are of great use to improve GST-based applications.

Crystalline GST

To reveal the structural features of crystalline GST alloys at atomic scale, experiments have been carried out using high-resolution scanning transmission electron microscopy (STEM), in which high angle annular dark field (HAADF) imaging and

energy dispersive X-ray spectroscopy (EDX) can be obtained.

Upon thermal annealing, amorphous GST undergoes crystallisation at 150°C, forming a polycrystalline metastable cubic phase. The metastable cubic phase has two sublattices: one fully occupied by Te atoms (anions), and the other randomly occupied by Ge / Sb atoms (cations), and vacancies. In cubic $Ge_2Sb_2Te_5$, vacancies take up 10% of the overall lattices [9, 10].

Figure 1.2 shows the HAADF images of cubic $Ge_2Sb_2Te_5$ obtained from different crystal orientations, including [001], [111], [110], and [211] [10]. The contrast of the HAADF images is roughly proportional to the square of the average atomic number (Z) of the atoms in one column: brighter spots in HAADF images correspond to heavier elements (i.e., larger averaged Z values). HAADF images taken along the [001] and [111] directions show no difference in the contrast of each atomic column, because all columns contain Ge, Sb, Te atoms, and vacancies, resulting in very similar average Z values for all spots in the HAADF images. When taken from the [211] and [110] directions, the two sublattices of the cubic phase can be well distinguished – the contrast of one sublattice is much brighter than the other one in the HAADF images. The former corresponds to the Te column whereas the latter is a mix of Ge, Sb, and vacancies [10].

The elemental distribution in crystalline GST can be better revealed in the EDX spectrum (Fig. 1.2). Along the [110] direction, Ge, Sb, and Te atoms are rendered by red, blue, and green spots, respectively. It is evident that Ge and Sb atoms occupy one sublattice in a random fashion whereas Te atoms occupy the other sublattice.

As the annealing temperature further increases to 180°C, the randomly distributed vacancies gradually become ordered and aggregate in some local regions of the sample – which typically occurs on the specific (111) planes. At this stage, the sample still exhibits non-metallic behaviour in electrical transport measurements [10]. The structure still remains in the cubic phase, and the overall degree of vacancy ordering is insufficient to reach extended electronic states [11]. As the annealing

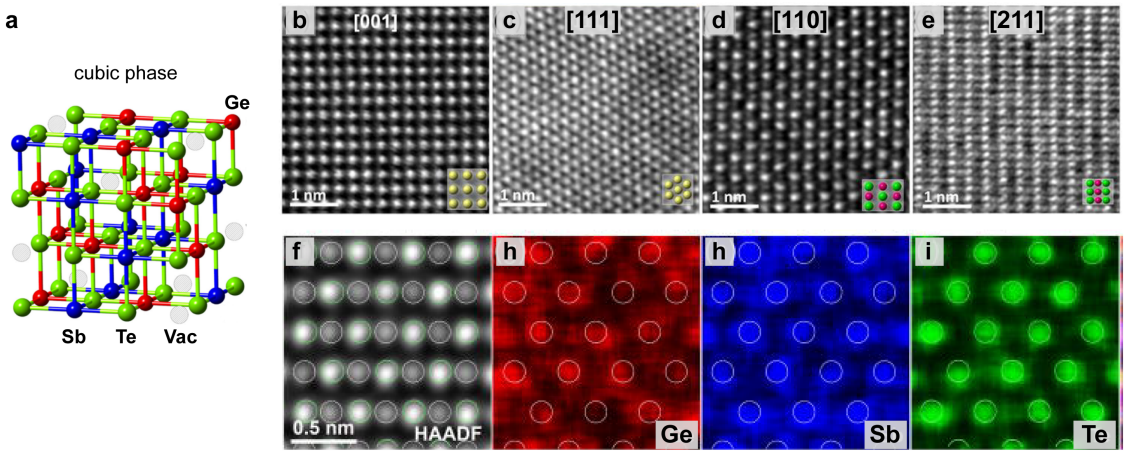


Figure 1.2: Atomistic models and STEM-HAADF images of rock-salt GST. (a) The cubic structural model of GST with Te (green spheres), Ge (red spheres), Sb (blue spheres), and vacancy (grey circles). HAADF images are taken along four crystal orientations: (b) [001], (c) [111], (d) [110], and (e) [211]. In the bottom right corner of each panel, possible atomic arrangements are depicted. (f) The HAADF image taken from the [110] direction. (h–i) the respective EDX chemical mapping for individual Ge (red), Sb (blue), and Te (green). The figure is reproduced from Ref. 10. Original figure published under the terms of the Creative Commons CC BY license.

temperature further increases, the number of ordered vacancies also increases. Such vacancy ordering proceeds until the temperature reaches 250°C, and the metastable cubic phase quickly transforms into a stable hexagonal phase (Fig. 1.3). For hexagonal $Ge_1Sb_2Te_4$, the crystal structure consists of repeating seven-layer blocks separated by van der Waals (vdW)-like gaps [12]. Different stacking sequences of the block unit have been proposed [13, 14], such as the Petrov model (Te-Ge-Te-Sb-Te-Sb-Te-Ge-Te), the Kooi model (Te-Sb-Te-Ge-Te-Ge-Te-Sb-Te), and the Matsunaga model (Te-Ge/Sb-Te-Ge/Sb-Te-Ge/Sb-Te-Ge/Sb-Te, where Ge/Sb represents mixed layers of Ge and Sb).

In HAADF images, the seven-layer blocks are clearly observed along the $[11\bar{2}0]$ direction (Fig. 1.3) [15]. The vdW-like gaps between the adjacent blocks are indicated by the arrows. Different contrasts are observed in the HAADF images, indicating different types of atomic columns. The brighter spots correspond to the heavier Te atoms and the slightly darker spots represent Ge, Sb atoms, or the mix of both. The EDX elemental analyses further distinguish different elemental columns

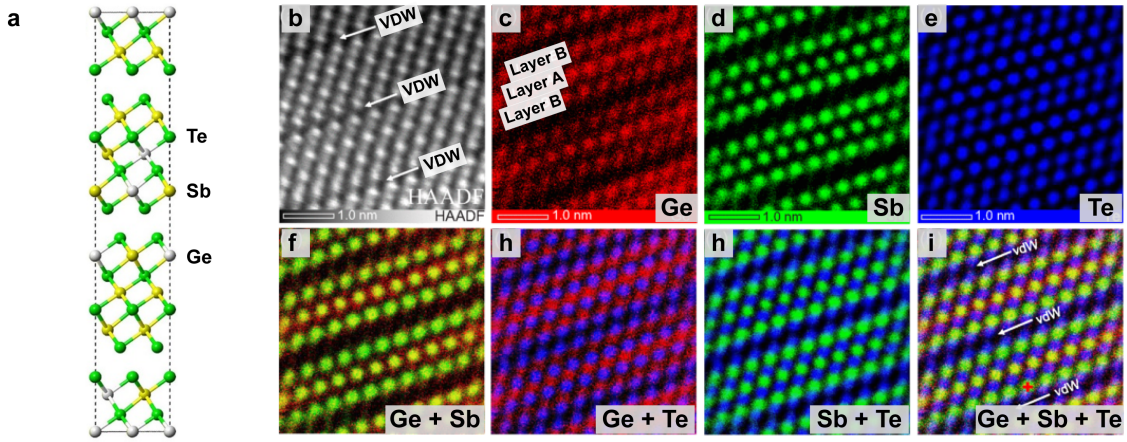


Figure 1.3: Atomistic models and STEM-HAADF images of trigonal GST. (a) The trigonal structural model of GST with Ge (white spheres), Sb (yellow spheres), and Te (green spheres). (b) The HAADF image taken along the [11$\bar{2}$0] direction. (c–i) EDX mappings for Ge (red), Sb (green), Te (blue), Ge+Sb, Ge+Te, Sb+Te, and Ge+Sb+Te atoms, respectively. The figure is taken with permission from Ref. 15. Copyright 2019 by the American Physical Society.

(Fig. 1.3) [15]. It is found that Ge/Sb sit in one sublattice and Te atoms are located in the other sublattice. The sublattice of Ge/Sb atoms can be further divided into two types: the central layer (layer A) away from the vdW-like gaps and the other two adjacent layers (layer B) close to the gaps, shown in Fig. 1.3. Notably, the intensity of the central Ge layer (marked as layer A in Fig. 1.3) is much higher than that of the two adjacent Ge layers near the vdW gaps (marked as layer B). The Sb layers show the opposite trend: the intensity of the central Sb layer is slightly weaker than that of the adjacent layers. Hence, these EDX results indicate that layer A contains more Ge (but less Sb) atoms than layer B, which is different from the proposed theoretical crystal models for hexagonal GST [13, 14]. Based on the experimentally characterised results, a clear atomic-scale picture of hexagonal GST can be obtained: Ge atoms tend to stay away from the vdW gaps, whereas Sb atoms tend to be closer to the vdW gaps [15]. Moreover, the structure of the hexagonal phase is more ordered than that of the cubic phase, explaining the three orders of magnitude difference in resistivity between the hexagonal and cubic phases of GST.

Amorphous GST

Unlike crystalline GST, the amorphous structure of GST is hard to fully characterise in experiments due to the loss of long-range structural ordering. In experiments, structural factors of amorphous GST alloys have been measured to study the overall medium-range order structural features [16]. To explore the atomic-scale information behind the measured structural factors, reverse Monte Carlo (RMC) simulations have been used to fit the X-ray diffraction data, which provides structural details of GeTe and $Ge_2Sb_2Te_5$ alloys [17]. Both liquid and amorphous structures of GST exist abundant primitive rings with different sizes (e.g., four-, five-, and six-membered rings), which reflect the medium-range structural complexity. These primitive rings continuously break and reform during the crystallisation process under thermal annealing, indicating a highly stochastic nature of the crystallisation.

Notably, in the crystalline GST, four-membered rings are predominant, together with a smaller number of six-membered rings. Given the similar structural building units, the quantified structural similarity between the amorphous and crystalline phases is of great importance to understand the crystallisation of PCMs, because the existence of four-membered rings in the amorphous or liquid phases provides a necessary condition for fast crystallization.

1.2.2 Aging and resistance drift

An important phenomenon of PCMs is aging, i.e., the structural relaxation of amorphous PCMs, which leads to considerable resistance drift in electronic devices. Aging is more prominent in amorphous PCMs compared to the crystalline phase. Resistance drift refers to the phenomenon where amorphous PCMs spontaneously undergo structural relaxation and exhibit an increase in electrical resistance over time. Taking $Ge_2Sb_2Te_5$ as an example, the resistivity (or the conductivity) of the amorphous phase consistently increases (decreases) over time (Fig. 1.4a) [18]. The temporal evolution of the resistance can be described by the equation $R(t) = R_0 \left(\frac{t}{t_0}\right)^{\nu}$, where

R_0 is the resistance at an arbitrary initial time t_0 shortly after the material has been amorphised, typically within a few hundred nanoseconds. At room temperature, the drift coefficient ν for the amorphous GST is measured to be 0.11217. In comparison, the drift coefficient for the crystalline phase is much lower, at only 0.0008 (Fig. 1.4a).

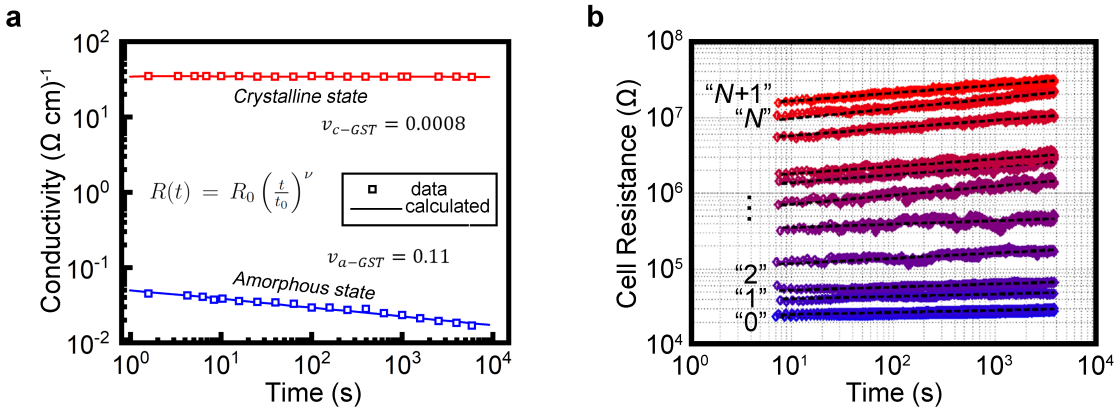


Figure 1.4: (a) Conductivity of amorphous GST and crystalline GST. An evident resistance drift is observed in amorphous GST. (b) Temporal fluctuations associated with the resistance values. The considerable resistance drift can lead to an error in multi-level storage applications using GST. Panel a is taken with permission from Ref. 18. Copyright 2009 AIP Publishing. Panel b is taken from Ref. 19. Original figure published under the terms of the Creative Commons CC BY license.

Note that the aging of amorphous GST is not a process of spontaneous recrystallisation. During the aging process, the amorphous phase undergoes spontaneous relaxation towards another disordered configuration with a lower energy. Therefore, the aging process observed in the amorphous GST occurs in an opposite direction of the crystallisation: the former tends to form a more stable glass whereas the latter leads to the crystalline phase. While the aging phenomenon of the amorphous phase does not directly impact the use of PCM-based binary memory/storage devices, it still poses a problem in multi-level memory/storage devices. Due to the aging effect, a cell initially set to state "N" may, after aging, exhibit resistance drift into the resistance window of state "$N+1$" [18, 19], likely leading to data errors and the unreliable operation of the device (Fig. 1.4b).

To understand the origin of resistance drift, a series of experimental characterisations were carried out on GST and its parent compound GeTe. It is found that the optical bandgap of GST increased monotonically in a logarithmic manner over time [20]. A similar electronic bandgap widening was also observed in amorphous GeTe [21]. Photothermal deflection spectroscopy (PDS), which is highly sensitive to subtle changes in the electronic structure and optical absorption, was utilised to monitor these changes over time. The PDS curves showed that the optical bandgap of amorphous GeTe gradually increased during the aging process [21].

In addition, modulated photocurrent experiments showed the distribution of defect states in the amorphous structure [22]: the density of deep defect states decreases after aging. According to the Poole–Frenkel effect, a reduction in defect state density increases the average distance between defects, thus increasing the energy barrier for electron transitions, which leads to an increase in electric resistance. However, for shallow defect states with an energy level less than 0.35 eV above the valence band, the defect density increases after aging [22].

Note that thermal heating can accelerate the aging process of the amorphous PCMs. It is found that the internal stress of amorphous GeTe decreased in the aging process [23]. Mechanical stress tests on 500-nm thick amorphous GeTe films showed that the as-deposited samples exhibit a compressive stress of 49 MPa. Upon heating to 50°C, the thin films underwent thermal volume expansion, which increases the compressive stress. However, as the structural relaxation of the amorphous phase occurred, the compressive stress kept on decreasing. Upon removal of thermal heating, the stress decreased again. The difference in the compressive stress before and after the thermal heating was around 16 MPa, indicating that structural relaxation can release the internal stress of amorphous PCMs.

Nevertheless, these experimental results require further understanding at the atomic level, including the atomic structure of defects, electronic states at defect sites, and the bandgap characteristics of the amorphous phase.

1.2.3 Volatile Ovonic threshold switching materials

In addition to the phase-change memory effect, some phase-change chalcogenides also exhibit another behaviour known as the Ovonic threshold switching (OTS) effect. As shown in Fig. 1.5a, when a voltage is applied to these phase-change chalcogenides, the current remains near zero at low voltages at the starting (i.e., a highly resistant state), which is regarded as the "OFF" state (i.e., a switched-off state). As the voltage keeps on increasing to a certain value, which is called the threshold voltage (V_{th}), there is a sudden increase in current, indicating the appearance of a highly conducting state. Sometimes, a negative differential resistance is observed at this stage. If the voltage continues to increase, the current also increases, and the materials remain in the conducting ON state. When the voltage is removed from the conducting selector materials, the materials quickly switch back to the initial OFF state with high resistance. In contrast, in phase-change memory materials, the resistance state remains upon removal of the external voltage, showing a non-volatile behaviour.

The OTS effect has been used in selectors for memory devices: when the applied voltage is below V_{th}, the device remains off; if the voltage exceeds V_{th}, the device switches to the ON state. A selector utilising the OTS effect is termed an OTS selector and the corresponding material is called an OTS material. When a voltage below the threshold voltage V_{th} is applied, only a minimal current, I_{OFF}, flows through the selector. However, when a read voltage exceeding V_{th} is applied, the selector maintains a large current, I_{ON}, sufficient to operate the target PCM cell. In practical devices, the OTS selector is stacked with the resistor (PCM) memory and such configuration is known as "1S1R" (1 selector + 1 resistor).

To obtain a high-performance OTS selector, several critical criteria have to be satisfied [24]. Firstly, to enable rapid data read and write operations, PCM cells must switch between ON and OFF states quickly under different voltage pulses. Hence, the operation time of OTS materials should be compatible with that of the PCMs.

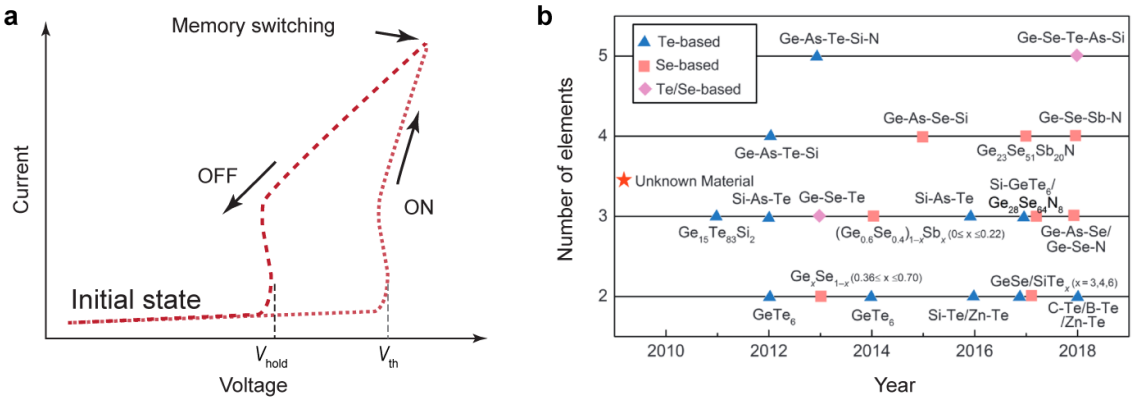


Figure 1.5: (a) Voltage–current curve of an OTS device. A volatile behaviour is observed. The figure is taken with permission from Ref. 24. (b) Different material systems used for selectors. Copyright 2020, the Authors. Published by Springer Nature. Permission conveyed through Copyright Clearance Center, Inc.

This indicates that the OTS should be switched between ON and OFF within 100 ns. Secondly, the ratio of the ON current (I_{ON}) to the OFF current (I_{OFF}) should ideally exceed 10^4, to enable high selectivity in OTS, leading to decreased power loss and increased density for devices. Thirdly, the OTS should be manufacturable using processes compatible with PCMs technology. It must also have good thermal stability, maintaining its non-linear behaviour after the back-end-of-line stage (i.e., high-temperature annealing) for PCM integration [25]. Moreover, the OTS must provide a current density exceeding 10 MA/cm^2 to facilitate the pulse-induced melt-quench amorphisation process for the switching of PCMs. Last but not least, since PCMs are expected to have an endurance of over 10^6 cycles, the selector must also demonstrate a similar operational lifetime.

To satisfy the aforementioned requirements for OTS materials, a range of candidates that exhibit the volatile OTS characteristics have been identified. In 1964, Noverthover and Pearson found the OTS behaviour in As–Te–I compounds [26]. Two years later, Ovshinsky reported the OTS phenomenon and proposed a potential device structure in a patent [27]. In 1968, Ovshinsky reported the OTS behaviour of the chalcogenide compound $Si_{12}Ge_{10}As_{30}Te_{48}$, and noted that removing arsenic from the compound led to non-volatile phase-change memory behaviour.

In recent years, numerous new OTS materials have been reported, as shown in Fig. 1.5, which can be categorised into three different types. The first type consists of Te-based alloys, and a widely reported example is Te-rich Ge-Te alloys [28, 29]. To improve the performance of this type of materials, various doping elements have been considered, e.g., silicon, nitrogen, boron, and carbon [30–33]. By doping, these materials exhibit fast switching speed. For example, $GeTe_6$ [29] and Ge–As–Te–Si–N alloys [32] can switch within 5 nanoseconds, and B–Te [34] and C–Te [35] exhibit even faster switches of less than 2 nanoseconds. The second type is Ge–Se-based alloys, attracting increasing attention due to their low I_{OFF} and good thermal stability [36–38]. Elemental dopants have also been considered in this type of materials, including N, As, Si, and Sb [37–39]. The third type contains both Te and Se, such as Ge–Se–Te–As–Si alloys [40].

However, as more and more dopants are considered in the OTS materials, which indeed improve the overall performance, it becomes more challenging to precisely control the stoichiometry of these multi-component compounds (e.g., involving more than five different elements). More importantly, under frequent external electric fields, ions may migrate towards the electrodes, which leads to severe phase segregation after extensive cycling and thus results in device failure.

More recently, a new Te-based OTS material has been proposed which used the spontaneous crystallisation of Te during the cooling process to achieve volatile switching behaviour. Due to a much more simplified composition, i.e., containing only one element, this elemental electric switch can avoid phase segregation and thus serve as a promising candidate for PCM-based memory devices. The study of phase transitions in Te as well as the Te-based selector is given in Chapter 4.

1.2.4 PCM-based memory devices

Given different purposes of applications, the cell structures of PCM-based devices are generally categorised into two main types: contact-minimised cells and volume-minimised cells (Fig. 1.6) [41]. The former controls the programming area by the

size of the bottom electrode, and the latter minimises the volume of PCMs within the cell, which are also called confined cells.

To achieve a storage density comparable to the maximum flash memory density at $4F^2$ (where F stands for the feature size), the confined cells are usually combined with a crossbar architecture, leading to one of the most promising structures for commercial devices (Fig. 1.6a), such as the Intel 3D cross-point storage and memory [8]. A crossbar contains five different layers that are alternatively stacked up. More specifically, main functional layers are sandwiched by three bit or word lines (i.e., each main functional layer sandwiched between two electrodes) (Fig. 1.6b). In the confined cell, the PCM memory unit is surrounded by dielectric materials, and hence the total volume of the PCM cell is fixed, leading to a constant atomic density.

When a specific confined cell needs to be operated, a bias voltage is applied between two corresponding electrodes. Different selection schemes have been used during operations, such as "1/2 V" scheme (that is, when the read voltage, V, is applied to the targeted phase-change memory cell, other lines in the array only remain biased at half the voltage, $1/2\ V$) [42]. However, when voltage is applied to the word line and bit line, cells that share the same row or column as the selected cell also experience some bias, creating unwanted current paths, leading to a phenomenon known as "leakage current". This leakage can cause several problems, including reduced resistance contrast between the logic states, increased power consumption, and unwanted operations of unselected units. To precisely control the ON/OFF of the memory cell, a switch unit is required [42–44], i.e., the Ovonic threshold switch mentioned above. Once the bias voltage reaches the threshold voltage (V_{th}), the memory cell will be switched on for read, write, or erase operations. The quality of the selector directly impacts the functionality of the memory.

In addition to the confined cells in a crossbar array, the other widely used device setup for PCM-based devices is the mushroom-type structure (Fig. 1.6c). In the mushroom-type device, the PCM layer is sandwiched between a top electrode and

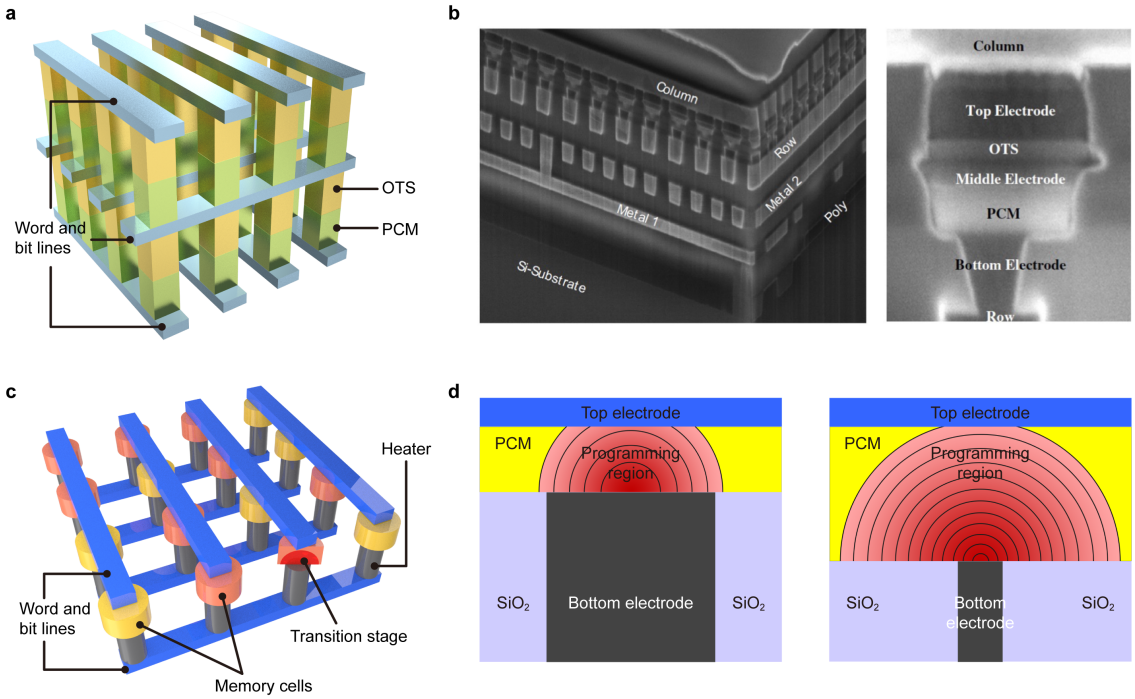


Figure 1.6: Two different types of PCM devices. (a) A sketch of 3D XPoint setup [8]. Confined cells are integrated in a crossbar array. Memory units (green) are connected by Ovonic threshold switching selectors (yellow) to prevent current leakage. (b) One layer of the crossbar array. (c) Sketch of an array consisting of mushroom-type devices. (d) Two-dimensional sketches of the mushroom-type device. The programming region is highlighted (in red). The resistance level depends on the ratio of the crystalline to the amorphous volume. Two different device configurations are shown which provide different intermediate resistance states (indicated by different semi-circles). Panel a is reproduced with permission from Ref. 1. Copyright 2019, Springer Nature Limited. Panel b is adapted with permission from Ref. 45. Copyright 2024 IOP Publishing. Panel c is reproduced with permission from Ref. 46. Copyright 2019, the American Association for the Advancement of Science. Panel d is taken with permission from Ref. 47. Copyright 2019, the American Association for the Advancement of Science.

a bottom electrode. Given a small radius of the bottom electrode, the PCM layer is only partially programmed (amorphised or crystallised), leading to an active region with a shape that resembles a "mushroom" (Fig. 1.6d). The mushroom-type cell is also called the "planar" structure, reflecting a large PCM unit. In contrast, the dimension of confined cells is usually smaller than that of mushroom-type cells and thus confined cells in a crossbar array can hardly be partially programmed. The small cell size and confined geometry suggest high-density binary memory applications. More importantly, the thermal disturbance (e.g., via heat dissipation)

between neighbour cells is greatly improved in the confined cell, also indicating its usefulness in commercial memory devices.

In fact, the resistance level in a mushroom-type PCM-based device depends on the ratio of the crystalline to the amorphous volume. Hence, the cell geometry can be designed to achieve different numbers of logical states. For example, using the same RESET programming parameters, a larger number of amorphous-crystalline volume configurations with distinguishable resistance values can be obtained if the thickness of the PCM layer increases and/or the diameter of the bottom electrode decreases (Fig. 1.6d) [47].

1.2.5 PCM-based in-memory computing devices

In addition to the PCM-based memory/storage applications, PCMs have also been used in advanced computing technology, e.g., brain-inspired computing chips [48–51], which aim to mimic the functioning of biological brains using electronic devices [52–57]. The core idea of brain-inspired computing is to enable in-memory computing, also known as neuromorphic computing, using arrays of memory units.

The conventional von Neumann computing architecture separates processing and memory units. For every operation, data are sent to the central processing unit (CPU) for processing and then transferred back to the memory units [1]. A hierarchy exists in different storage and memory devices, including the cache (SRAM) embedded in the CPU, the memory (DRAM) that frequently exchanges data with the CPU, and non-volatile storage such as solid-state drives (SSD) and mechanical hard drives (HDD). The transfer of large amounts of data between these devices restricts further improvements in computational power and results in approximately 40% of ineffective energy consumption [57].

To eliminate heavy data transfer and enable large-scale parallel computing and storage tasks, in-memory computing unifies the processing and storage of data in the same functional unit, in the absence of frequent data shuffling. The underlying working principle of in-memory computing is to utilise the multi-valued character-

istics of memory units and the intrinsic physical properties of materials (such as Kirchhoff's law and Ohm's law) to perform direct computations. For example, a vector-matrix multiplication can be realised by simply using a crossbar array and memory cells as mentioned in Section 1.2.4.

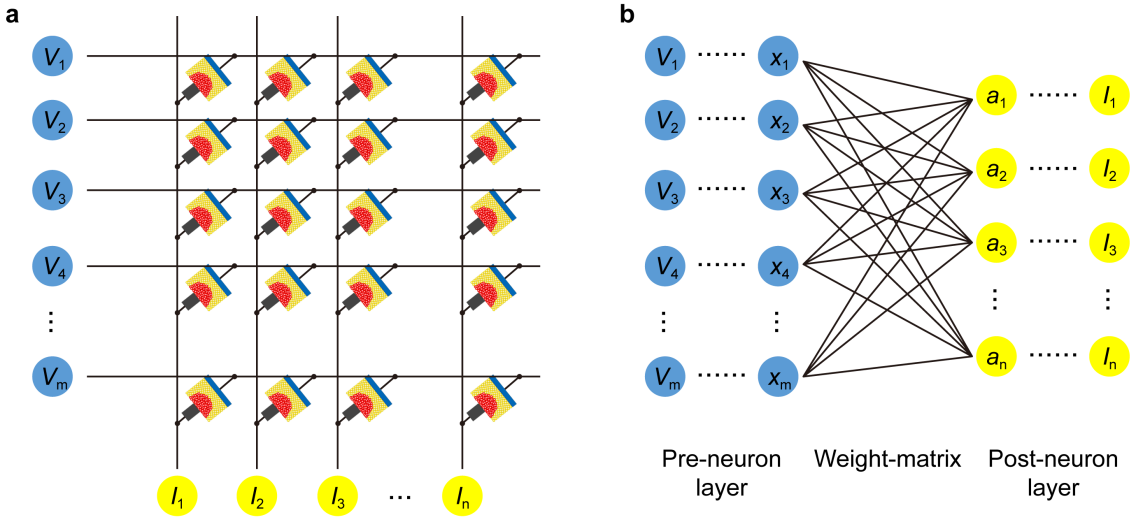


Figure 1.7: Schematic diagrams illustrating the principle of matrix multiplication using a crossbar array. (a) A crossbar array of PCM-based devices. (b) A vector-matrix multiplication is realised using the input voltage values and output current values of the crossbar array. The matrix elements are represented by the conductance of the PCM cells. The schematic sketches were drawn after Ref. 58.

As shown in Fig. 1.7a, voltages are applied to the bit lines (denoted as V_1, V_2, V_3, ...). According to Ohm's law, the current through a memory cell is equal to the voltage across the memory cell multiplied by its conductance. Following Kirchhoff's current law, the current on each word line is the sum of the currents through each node (denoted as I_1, I_2, I_3, ...). If the input voltage values on each bit line are considered as the input variables in a neural network (denoted as x_1, x_2, x_3, ...), and the measured output current values as the output variables of the neural network, then with fixed weights at each neural node represented by the conductance of memory cells, the desired output values of the neural network (denoted as a_1, a_2, a_3, ...) can be directly obtained through a single circuit operation (Fig. 1.7b).

Hence, the mathematical process of vector multiplication can be realized through the basic operation of applying voltage and measuring current, bypassing the com-

plex and separate data processing and storage operations in traditional binary von Neumann computers. Given that the state-of-the-art machine learning technologies involve abundant and frequent vector multiplication operations, in-memory computing aligns perfectly with ML operations, which achieve significant improvements in computational, energy, and volumetric efficiency.

In fact, the idea of using non-volatile phase-change materials for neuromorphic computing applications was proposed decades ago [52]. However, due to significant resistance drift and high resistance noise (Fig. 1.4) [19], new material systems are required to enable robust in-memory computing tasks. The application of PCMs in neuromorphic computing and in-memory computing has gradually been realised in recent years [48–51]. By using multiple short programming pulses to achieve cumulative crystallization (cumulative SET) and gradual amorphisation (iterative RESET) of PCMs, stable multi-level logic states can be obtained using the mushroom-type devices in a crossbar array (Fig. 1.6c). A detailed discussion regarding these two key operations will be discussed in Section 1.3.3.

In addition, benefiting from recent advancements in silicon waveguide technology, PCMs have also been successfully used in photonic devices [59, 60], which leads to photonic non-volatile memory [61–65], photonic neuromorphic computing [66–71], and flexible display technology with nanoscale resolution (nano-pixel display) [72–75]. However, this thesis focuses on the PCM-based electronics and the study of PCM-based photonics will be expected in the future.

1.3 Working mechanisms of phase-change materials

Unlike the conventional silicon-based storage and memory (e.g., SSD, DRAM, and SRAM) whose working mechanisms are largely dependent on charge levels, phase-change random access memory (PCRAMs) use the large electrical or optical contrast between the amorphous state and crystalline state of PCMs. The amorphous state

has a high electric resistance but a low optical reflectivity whereas the crystalline state has a low electric resistance but a high optical reflectivity. Such different properties of PCMs stem from differences in structural disorder, carrier concentration and bonding mechanisms [11, 21, 76–78].

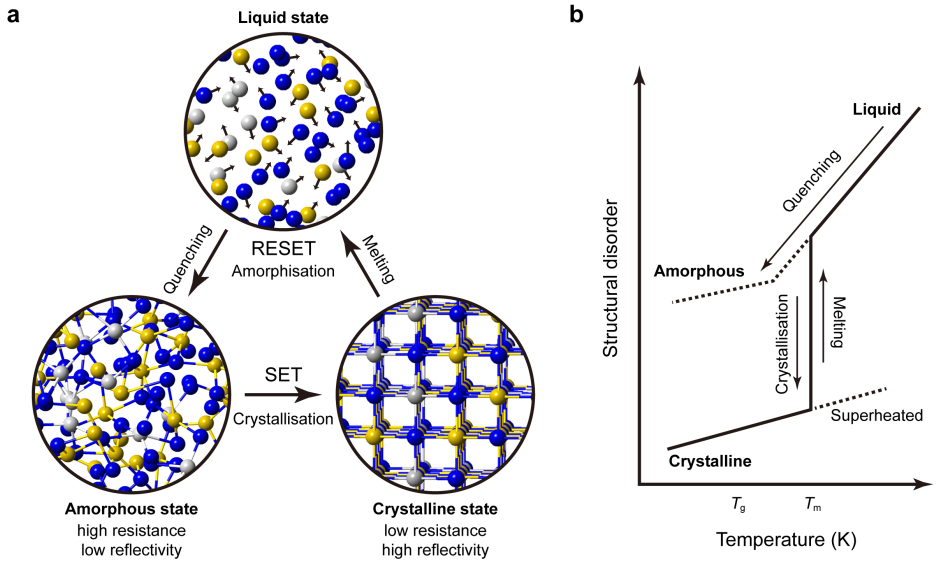


Figure 1.8: The working principle of PCMs for memory applications. (a) Relationships among different phases of GST alloys. Ge, Sb, and Te atoms are rendered as white, yellow, and blue balls, respectively. The amorphous and crystalline states are characterized by high resistance/low reflectivity and low resistance/high reflectivity, respectively. (b) Thermal paths in PCMs. Heating the crystalline states up to the melting point leads to the melting of GST. The fast cooling of the melted liquid state results in a supercooled amorphous phase. The annealing at the crystallisation temperature gives rise to a fast crystallisation. Panel a is adapted with permission from Ref. 6. Copyright 2007, Springer Nature Limited. Panel b is taken from Ref. 79. Copyright 2019, the Authors. Published by Springer Nature.

The fast and reversible phase transitions between the two states can be triggered by thermal heating (i.e., Joule heating): electric pulses are used in PCM-based electric devices and laser pulses are implemented in PCM-based optical devices. The amorphous and crystalline states are both thermally stable at ambient temperatures, leading to good data retention ability which is required for nonvolatile phase-change applications. In phase-change devices, the write (SET) operation is accomplished by a crystallisation process at elevated temperatures (500–600 K); and the erase (RESET) operation is achieved via an amorphisation process via a melt-quench process at temperatures above the melting point (Fig. 1.8). The amorphisation, i.e.,

a structural disordering process, can be as fast as several hundred picoseconds while the crystallisation process, e.g., in GST alloys, usually takes tens of to hundreds of nanoseconds, which becomes the bottleneck for further increasing the writing speed of PCMs.

1.3.1 Amorphisation of PCMs

One of the most common approaches for turning a crystal into a disordered phase is high-temperature thermal annealing. By heating the crystal to a temperature above its melting point, the crystal melts into a liquid. Rapidly cooling the disordered liquid down can prevent the system from having sufficient time to crystallise. During the cooling, the atoms "freeze" into a disordered configuration (Fig. 1.8). This is known as the "melt-quench" amorphisation process, which has been used to create amorphous forms of different materials. If the melt-quench process is carried out on a thin film, the whole sample is then annealed and then cooled down. Hence, this melt-quench corresponds to a isothermal heating process where the temperature of different parts of the solid is nearly constant and no evident temperature gradient is observed.

In phase-change memory devices, heating is achieved through Joule heating generated by passing an electric current through the device. The RESET process facilitates the phase transition from a crystalline to an amorphous state. The electric pulse, added on the low-resistance crystalline PCMs, generates significant Joule heat, thereby heating the material up quickly. The heating process is very swift, due to a short pulse duration, and the liquid rapidly undergoes a cooling process, dissipating heat to the surrounding environment and forming an amorphous state. Unlike the isothermal heating in the melt-quench process of a thin film, the amorphisation via an electric pulse creates a considerable temperature gradient, corresponding to a non-isothermal heating.

Note that the energy consumption of the amorphisation process contributes significantly to the total energy consumption of the device. The melting behaviour

of PCMs depends not only on the intrinsic properties of PCMs, e.g., the melting point, but also largely on the overall thermal design of the device. Reducing the heat dissipation to the surrounding environment during the heating process can greatly reduce the energy consumption of the device. In poorly optimised devices, most of the heating energy can be wasted due to the heat dissipation.

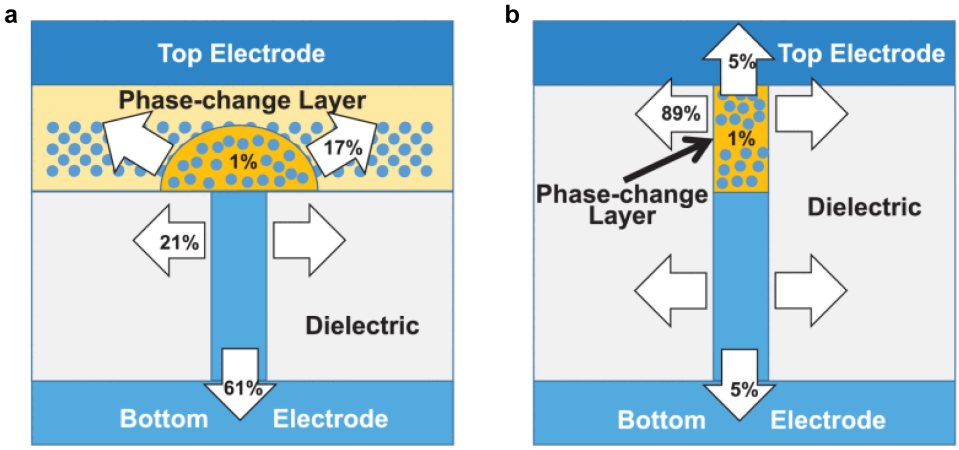


Figure 1.9: Schematic of a typical (a) mushroom and (b) confined cell using PCMs. Arrows represent heat loss paths in the phase-change layer during RESET operations in the respective cells. The heating follows a non-isothermal fashion, where a considerable temperature gradient exists. The figure is adapted with permission from Ref. 80. Copyright 2017, IEEE.

As shown in Fig. 1.9, the main energy loss in the device occurs through the heat transfer to the bottom electrode and the dielectric layer. Optimising the electrode and dielectric materials to increase their intrinsic thermal resistance and their interfacial thermal resistance will help reduce energy loss during heating. For example, carbon materials can reduce the RESET current by 75% and 40%, respectively [81, 82], largely reducing the energy consumption of RESET operations. In addition, multilayer dielectric layers can also increase the thermal resistance of the dielectric layer, thus reducing energy consumption [83]. In addition, miniaturising the dimensions of the device can further increase the thermal resistance [84, 85].

Despite the largely reduced energy consumption via the increase of thermal resistance, the cooling process will also be affected: when the thermal resistance increases, the cooling rate decreases accordingly. The slow cooling might lead to a fast

crystallisation inside PCMs, thereby resulting in a failure of the RESET process. Hence, it is crucial to keep a good balance between the increased thermal resistance and a sufficiently high cooling rate to obtain the desired amorphous state after the melt-quench process.

1.3.2 Crystallisation of PCMs

Crystallisation plays a significant role in operating PCM-based devices. The crystallisation of PCMs is dominated by nucleation and growth, and hence, two types of PCMs are identified: the nucleation-dominated and growth-dominated PCMs.

In nucleation-dominated PCMs, the starting of crystallisation is usually signalled by the formation of nanometre-scale crystalline seeds (i.e., "nuclei"). As shown in Fig. 1.10, small crystalline grains gradually appear within the amorphous region at the crystallisation temperature. With further annealing, these small grains gradually grow and eventually crystallise to form polycrystalline regions with varied crystal orientations. GST alloys are a typical example of a material system where both nucleation and growth predominantly exist. Using an electron microscope, it is found that the nucleation regions exhibit a different contrast with respect to the surrounding areas (Fig. 1.10), indicating the formation of a polycrystalline domain. Moreover, the grain size largely depends on the crystallisation temperature, with dimensions ranging from a few nanometres to tens of nanometres.

In contrast to nucleation, crystal growth appears at the interface between the crystalline and amorphous regions. As illustrated in the upper right of Fig. 1.10, at high temperatures, the crystallisation at the amorphous-crystalline interface proceeds rapidly. During the growth process, no stable crystalline grains form, resulting in the final structure being a single crystal rather than a polycrystal. Silver (Ag) and Indium (In) co-doped Sb_2Te (AIST) alloys (e.g., $Ag_4In_3Sb_{67}Te_{26}$) are a classic type of growth-dominated PCMs. As shown in the lower right of Fig. 1.10, the contrast of the recrystallised region is indistinguishable in comparison to the original crystalline matrix, indicating the formation of a single crystal.

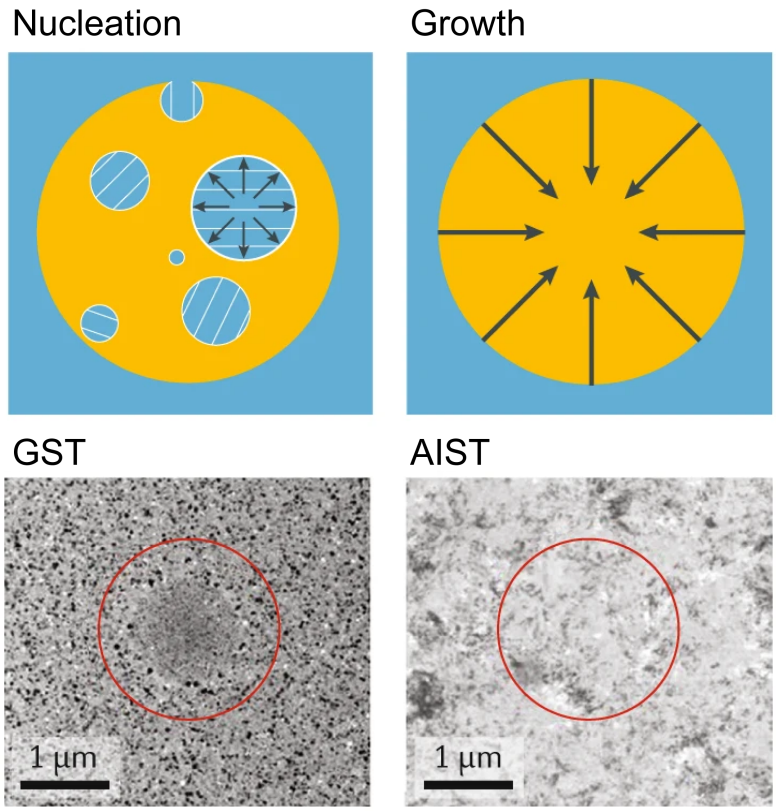


Figure 1.10: Comparison of nucleation- (e.g., GST alloys) and growth-dominated (e.g., $Ag_4In_3Sb_{67}Te_{26}$ (AIST)) PCMs. The transmission electron microscope (TEM) images show the crystallised phases upon thermal heating (marked by the red circles). In nucleation-dominated GST, the recrystallised region displays a clear contrast with respect to its crystalline surrounding owing to the difference in grain size and orientation. By contrast, growth-dominated AIST exhibits no contrast, indicating a smooth crystal growth from the boundary, yielding a single-crystalline state. The figure is adapted with permission from Ref. 1. Copyright 2019, Springer Nature Limited.

According to the classical nucleation theory, the difference of Gibbs free energy between the crystalline and amorphous states is the primary driving force of nucleation. However, the crystallisation rate also depends on the atomic mobility. At lower temperatures, although the nucleation rate is high, the low temperature results in very limited atomic mobility, thereby impeding the growth process. In contrast, at very high temperatures, frequent atomic diffusion is observed; however, the nucleation events are rarely seen.

Nucleation

According to classical nucleation theory, homogeneous nucleation is a stochastic process induced by thermal fluctuations. The nucleation consists of two competing effects: the bulk free energy and the interfacial free energy, which together determine the height of the free energy barrier for nucleation. The former is a volume effect, whereas the latter is dependent on the surface of the nucleus. The free energy of the system arises, due to the interface penalty, when atoms are added to form subcritical nuclei. As these small crystalline seeds grow up to the critical-sized nuclei, the free energy starts to decrease. The formation of a critical nucleus requires a long and stochastic incubation process. The larger the bulk free energy (or the smaller the interfacial free energy), the smaller the critical nucleus size, and vice versa.

Since both the driving force and the interfacial energy are temperature-dependent, the size of the critical nucleus varies with temperature. Based on the classic nucleation theory, the nucleation rate exhibits a maximum at a given temperature, which is determined by two factors competing with each other: with decreasing temperature, the bulk free energy (i.e., the driving force of nucleation) increases (reducing the required size of the critical nucleus), but the atomic diffusion rate also decreases (reducing the rate of atomic attachment to the nuclei). However, due to the non-isothermal heating conditions and non-equilibrium dynamics in phase-change devices, classical theory cannot fully explain the crystallisation of PCMs. Hence, it is necessary to delve into the atomic scale, exploring the structural and dynamic evolution during nucleation.

Growth

Unlike nucleation-dominated GST alloys that can undergo nucleation very quickly, growth-dominated materials usually have very long nucleation. In 2009, researchers at IBM investigated the nucleation process of AIST alloys [86]. They found that AIST has significant difficulty in nucleating, even at higher temperatures, making it challenging to form critical crystalline seeds. Researchers used laser heating to

characterise the crystallisation process based on the changes of optical reflectivity of AIST [86]. They found that the nucleation required for amorphous AIST alloys is extremely long, exceeding 20 microseconds [86]. This incubation time is several orders of magnitude longer than the nucleation time required for GST alloys. In-situ TEM experiments were also carried out to characterise the crystal growth of AIST at low temperatures. It was found that the recrystallised region has no contrast as compared to the surrounding crystalline matrix that previously existed. This, therefore, indicates that AIST formed a single crystal via crystal growth.

As the temperature increases, the growth rate of AIST increases significantly. At higher temperatures, the growth speed of the material becomes too fast to be measured in TEM experiments. An alternative method involving a dual-laser setup was employed to determine the growth rate of AIST by measuring the changes of optical reflectance over time [87]. Note that a decrease of optical reflectance indicates the formation of the amorphous phase, while a gradual increase of reflectance indicates a recrystallisation process. By calculating the change rate of the optical reflectance over time at a given temperature, the growth rate of AIST could be determined at that temperature [87]. Based on the experiments, it is found that the growth rate of AIST alloy varies drastically with temperature [87]. As the temperature increases from 418 K to 553 K, the growth rate of AIST can increase from 100 nm s^{-1} to over 3 m s^{-1}, covering almost eight orders of magnitude [87].

Note that the crystal growth process is also vital for nucleation-dominated materials (e.g., GST alloys). As critical nuclei form after the incubation, a fast crystal growth is observed, starting from the interfaces between the crystalline seeds and the surrounding amorphous phases. To better understand the intrinsic mechanism of the crystal growth and the difference between nucleation- and growth-dominated materials, atomic-scale observations will be required. A detailed introduction of theoretical studies of crystallisation of PCMs will be given in Section 1.4.

1.3.3 Cumulative SET and iterative RESET

In addition to binary storage and memory, multilevel storage has been demonstrated to be feasible using PCM-based electronic devices [1]. As discussed in Section 1.2.5, two important working modes are required for PCM-based in-memory computing devices, i.e., the cumulative SET and iterative RESET operations. As shown in Fig. 1.11a, instead of amorphising the memory cell using one single RESET pulse, a series of high-level and narrow voltages or laser pulses with different amplitudes are applied to the crystalline PCMs. These pulses heat up the programming region of the crystalline memory cell, and the different amplitudes will lead to different amorphised volumes after each pulse (Fig. 1.11b). Since the resistance of the cell changes as a function of the ratio of crystalline-to-amorphous volume. Hence, different resistance levels can be obtained in the different RESET states. Such a RESET programming scheme on the crystalline PCMs is called iterative RESET.

In contrast, during the SET process, instead of using one long single SET pulse (over several to tens of nanoseconds) to fully crystallise the amorphous region, a train of weak, short, but constant SET pulses can be applied sequentially to obtain a gradual crystallisation in a cumulative way (Fig. 1.11c). Due to the constant pulse amplitude, the programming region of each pulse remains the same, whereas the effective amorphous volume (which is determined by the ratio of crystalline to amorphous volume) changes in a random fashion, affected by the stochasticity in nucleation and crystal growth (Fig. 1.11d). Such a SET programming scheme on the amorphous part of PCMs is termed cumulative SET.

As for in-memory computing applications, both iterative RESET and cumulative SET operations are used to obtain different resistance levels in the memory cells, which are then used to precisely control the different weight values in the neural network. In addition, such gradual crystallisation and amorphisation processes mimic the biological synapse, which can be used to emulate the synaptic plasticity in the brain. However, both iterative RESET and cumulative SET modes suffer from

considerable variation, which hinders the reliable and robust in-memory computing operations. For example, iterative RESET suffers from spontaneous structural relaxation of amorphous PCMs at ambient temperatures [21], leading to considerable resistance drift [88]. Moreover, due to the stochasticity of crystallisation, the cell resistance changes in a non-linear way. Considerable fluctuations of cell resistance are found from pulse to pulse and from cycle to cycle, causing large variability.

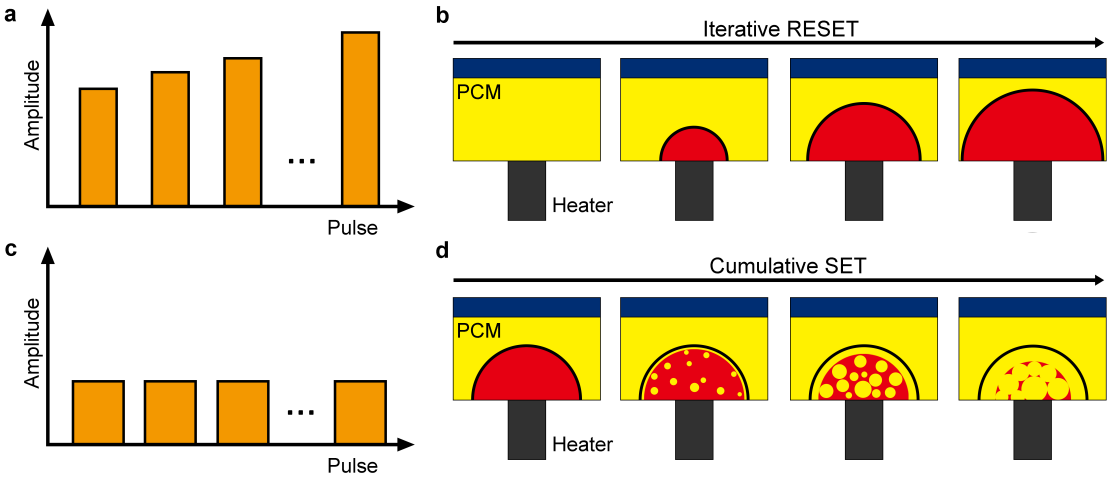


Figure 1.11: Two different working modes of PCM-based in-memory computing, including (a, b) iterative RESET operation and (c, d) cumulative SET operation. The amorphous and crystalline parts of PCMs are rendered in yellow and red, respectively. The figure is taken with permission from Ref. 79. Copyright 2019, the Authors. Published by Springer Nature.

1.4 Ab initio atomistic simulations of phase-change materials

1.4.1 Ab initio molecular dynamics simulations

Unlike the experimental studies on crystalline PCMs, which clearly reveal the atomic arrangement of crystalline GST and random distribution of atomic vacancies, amorphous structures of PCMs are hard to be fully characterised in experiments, due to the loss of medium-range structural order. To obtain a clear figure of amorphous PCMs, atomistic simulations play an important role. In this section, different widely used atomistic simulation methods will be discussed, showcasing how theoretical

modelling can help to understand complex structural features and the dynamical process of PCMs.

Structural properties of amorphous GST

To understand the structural features of amorphous PCMs at the atomic scale, melt-quench simulations were carried out using AIMD to create structural models [89–92]. For example, the structure factor of $Ge_2Sb_2Te_5$ obtained from X-ray diffraction experiments was compared with that obtained from AIMD simulations [89] – a good agreement is observed between the experimental and theoretical data [89], which suggests the feasibility of AIMD simulations in the structural characterisation of amorphous GST at the atomic scale.

Coordination analysis of the amorphous GST model reveals that the primary coordination numbers of cations (Ge and Sb) and anions (Te) are 4 and 3, respectively, which are lower than the ideal octahedral coordination number of 6, due to the presence of atomic vacancies [89]. However, the bond angles remain close to 90° and 180°, similar to the metastable cubic phase, indicating a large presence of distorted octahedral configurations formed by undercoordinated Te and Ge/Sb atoms.

Local structural analyses showed the presence of four-coordinated Ge atoms in a tetrahedral configuration, in addition to the abundant distorted octahedral motifs. A significant number of short homopolar Ge-Ge and Ge-Sb bonds exist, indicating the formation of tetrahedral motifs [21]. About one-third of the Ge atoms in the amorphous GST exhibit tetrahedral coordination, with the average bond length in the tetrahedral configuration (2.73 Å) shorter than that in the distorted octahedral configuration (2.82 Å) [89]. These simulation results are consistent with models obtained based on the extended X-ray absorption fine structure (EXAFS) and X-ray absorption near edge structure (XANES) data [89].

Aging mechanisms in GeTe

To explore the microscopic mechanism of aging in amorphous PCMs, AIMD simulations were carried out to provide atomic-scale structural information. Raty et

al. utilised AIMD simulations and an elemental substitution scheme to generate a series of amorphous GeTe models with varied proportions of Ge-Ge homopolar bonds [21]. Specifically, they created two different amorphous structures of GeTe: Ge atoms were substituted for Sn in a melt-quenched amorphous SnTe model, and Te atoms replaced Se in a melt-quenched amorphous GeSe model. After manual substitution of cations or anions, the structural models were annealed at room temperature for thermal annealing. The resultant structures thus have different ratios of Ge-Ge homopolar bonds due to the distinct bond proportions originally present in SnTe and GeSe amorphous models. The proportion of tetrahedrally coordinated Ge atoms was found to correlate approximately linearly with the proportion of Ge-Ge homopolar bonds (Fig. 1.12a–b). As the proportion of tetrahedrally coordinated Ge atoms decreased, the overall system energy also decreased significantly.

The results showed that the aging process involved a gradual reduction of tetrahedrally coordinated Ge atoms in amorphous GST (Fig. 1.12a). As the proportion of tetrahedrally coordinated Ge atoms decreased, the internal stress also decreased. Such theoretical results aligned with the experimental evidence that the aging process in GeTe is accompanied by stress relief [23]. A higher proportion of tetrahedrally coordinated Ge atoms corresponded to a larger internal stress, indicating that the presence of tetrahedrally coordinated Ge atoms and the Ge-Ge homopolar bonds destabilised the amorphous structure of GeTe. Note that these bonds and the tetrahedrally coordinated Ge atoms are kinetically stabilised under rapid quenching conditions but are thermodynamically unfavourable.

In addition to the reduction of Ge-Ge homopolar bonds and tetrahedrally coordinated Ge atoms, another structural key to aging is the enhanced Peierls distortion (Fig. 1.12d–e). Peierls distortion refers to the formation of pairs of aligned long and short bonds with a bond angle of approximately 180°, which is widely observed in crystalline PCMs (e.g., trigonal GST structures). The degree of Peierls distortion can be assessed by calculating the angular-limited three-body correlation (ALTBC)

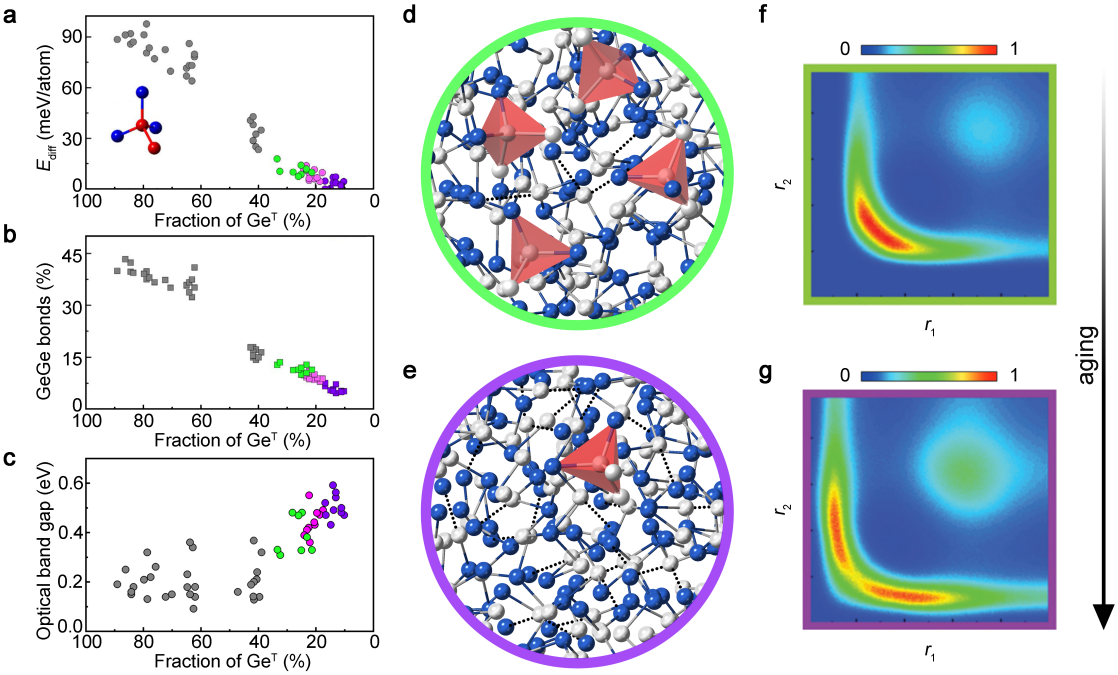


Figure 1.12: Aging of amorphous GeTe. Different amorphous models of GeTe were generated via the melt-quench AIMD simulations (green) or using elemental substitution based on amorphous SnTe (violet), GeSe (light magenta) and SiTe (grey). As the fraction of tetrahedrally coordinated Ge atoms decreases, (a) the energy difference decreases, (b) the fraction of homopolar GeGe bonds decreases, and (c) the optical bandgap increases. (d, e) The schematic atomic configurations of amorphous GeTe before aging and after aging are shown. (f, g) Sketches show the difference of normalised angular-limited three-body correlation (ALTBC) function between the pristine amorphous GeTe (before aging) and the aged amorphous GeTe (after aging). The black arrow indicates the direction of aging. Panel a is adapted with permission from Ref. 21. Copyright 2015, Springer Nature Limited. Panel b is adapted with permission from Ref. 93. Copyright 2020, Royal Society of Chemistry.

function [21]. As shown in Fig. 1.12f–g, the pristine structural model contains many pairs of bonds of the same or similar length, while after aging, the model features long and short bonds, indicating the enhanced Peierls distortion upon cooling.

The onset of Peierls distortion is also associated with the increased band gap (Fig. 1.12c), which increases the energy barrier of carrier excitation and thus reduces the electric conductivity, leading to resistance drift. After aging, it is found that the band gap of amorphous GeTe is widened, and the mid-gap defect states disappear [21]. These changes of electronic and optical properties were suggested to result from the structural changes during the aging process.

Simulating melting of PCMs

Given randomly distributed vacancies in the crystalline states, experimental characterisation is difficult to capture the atomic-scale picosecond timescale melting process in GST. In fact, the amorphisation process in PCM-based devices is extremely rapid: it requires only hundreds of picoseconds, which might be several orders of magnitude faster than the crystallisation (SET operation) [25]. Hence, AIMD simulations become feasible to access the microscopic evolution during the melting process, which can help to elucidate the underlying mechanisms.

A previous study, using DFT calculations, studied the changes of the electronic properties during the amorphisation of PCMs, which results from the local structural distortions and leads to an opening of the bandgap [94]. Sun et al. used AIMD simulations to study the melting behaviour of $Ge_2Sb_2Te_5$, and attributed the fast and reversible phase transitions to the formation of two-dimensional linear or tangled clusters [95]. These clusters vibrated around their initial configurations (i.e., the equilibrium positions) only in the perpendicular direction at the temperature (e.g., 800 K), slightly below the melting point.

Note that GST alloys have two crystalline phases: a metastable cubic structure and a hexagonal structure. When an amorphous GST thin film is annealed, it first crystallises into the cubic phase at 140°C and then transforms into the hexagonal structure upon heating to 310°C [96]. Further increasing the temperature will lead to the melting of hexagonal GST at 623°C (the melting point) [26]. In cubic $Ge_2Sb_2Te_5$, the anion sites are occupied by Te, while the cation sites are occupied by Ge and Sb, with 20% of cation sites being vacancies. These vacancies were suggested to play a crucial role in the structural evolution and physical properties of PCM, such as the metal-insulator transition [11]. The increased vacancy layers can reduce lattice thermal conductivity and improve the thermoelectric figure of merit in GeTe-rich GST, thus enabling tunable thermoelectric properties [97].

Given a considerable number of vacancies in cubic GST, it has been suggested,

based on AIMD simulations, that the melting process in the cubic phase of GST initiates from the vacancy layers and gradually propagates to the adjacent regions [98]. The rapid melting observed around the vacancy layers was attributed to the displacement of Ge atoms into the adjacent vacancy layers [99].

Although AIMD simulations are of great use to studying the microscopic mechanisms of the melting in GST, large structural models are required to better capture the behaviour of vacancy migration at elevated temperatures, e.g., the clustering of vacancies and the formation of large voids. Structural models containing only ≈200 atoms might be considerably affected by the finite-size effects. Larger simulation sizes should be considered which might go beyond the capacity of AIMD.

Atomistic simulations of nucleation

As mentioned in Section 1.3.2, nucleation requires a very long incubation due to its largely stochastic nature. A microscopic point of view at the atomic scale is of great importance to better understand this process. A combination of X-ray diffraction experiments and reverse Monte Carlo simulations has revealed the structural origin of the fast nucleation in GST alloys – the structural similarity between amorphous and crystalline GST [17]. Both phases have abundant four-membered rings and octahedral-like motifs, which accelerate the incubation process.

These results are also supported by AIMD simulations: a significant number of four-membered rings are found in both amorphous and liquid GST alloys [91, 100]. The existence of these four-membered rings is a critical factor for rapid crystallisation into the cubic phase. More importantly, it is found that four-membered rings took up the highest proportion of ring types in both the liquid and amorphous phases of GST alloys [91]. Most of these four-membered rings are ABAB type, where "A" represents cation (Ge/Sb) atoms and "B" represents anion (Te) atoms. The formation of ABAB four-membered rings is crucial during the nucleation process because they are also the smallest structural units of the cubic phase in GST alloys. The greater the number of ABAB four-membered rings is in the amorphous and

liquid structures, the more similar these structures are to the crystalline phase.

To further demonstrate the nucleation process of GST alloys at the atomic scale, AIMD simulations were carried out to simulate the nucleation and crystallisation process of GST alloys. Simulation boxes of no more than several hundred atoms were used, which were accessible with AIMD simulations at an acceptable computational cost. For example, the first crystallisation simulation of GST was based on a 72-atom structural model [100]. However, this model is largely limited by the finite-size effect, which showed a very fast crystallisation during the melt-quench process and no incubation process was observed. For larger simulation cells (containing more than 100 atoms), the finite-size effects can be reduced [101–107]. A 0.5-ns AIMD simulation was carried out using a structural model of GST containing 460 atoms [105]. Such simulations were obtained in several months of real-world time using parallel computing clusters.

Based on the AIMD simulations of 180 atoms, the evolution of local atomic motifs in the nucleation process was studied. As shown in Fig. 1.13a–d, the nucleation process in GST can be roughly divided into four different stages. At the first stage, abundant free-standing four-membered rings were found. Some of these four-membered rings started to connect, forming cubic-like clusters at the second stage. These cubic clusters gradually expanded as additional four-membered rings aggregated. When it comes to the third stage, the growing cubic clusters reached the critical size, thus becoming crystalline nuclei. The formation of these crystalline nuclei signals the end of the incubation process and the starting of a fast crystallisation. During this stage, a fast increase in the number of crystalline atoms, four-membered rings, and cubic motifs was observed. Lastly, the structural model underwent a fast crystallisation into a defective cubic phase.

Based on these AIMD simulations of the incubation process, quantitative structural analyses were carried out [101] (Fig. 1.13e). During the incubation, the four-membered rings are all rather isolated in the disordered matrix and no well-defined

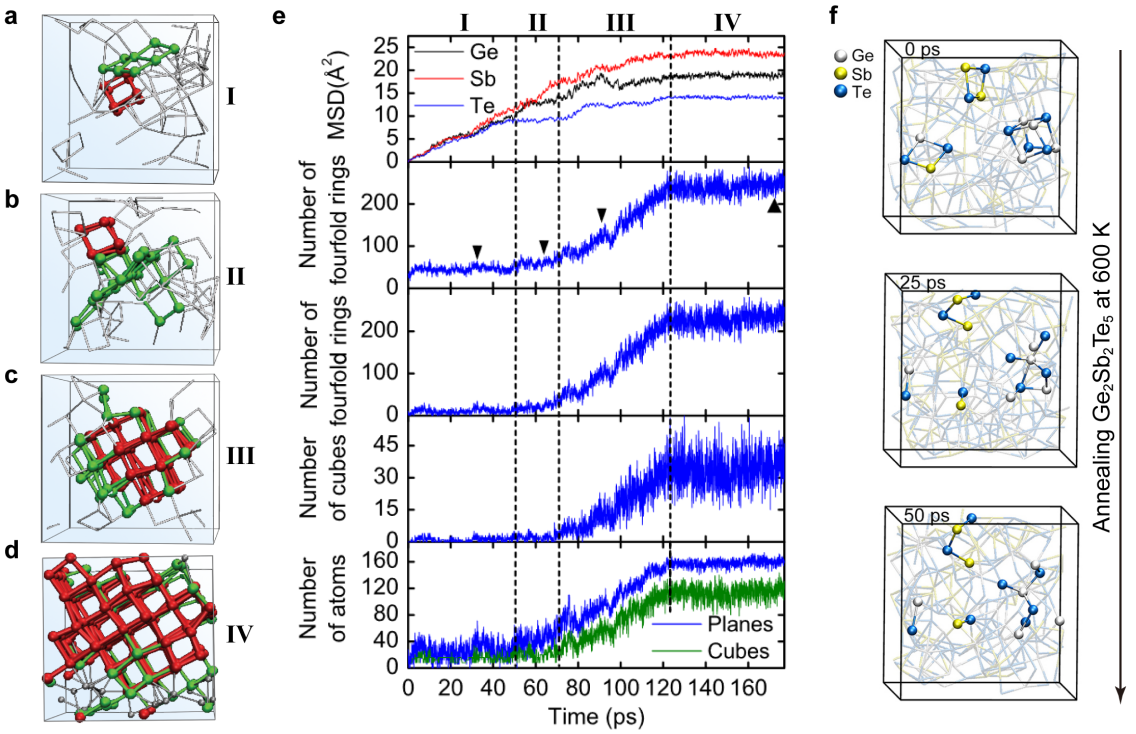


Figure 1.13: (a–d) Four different stages of the nucleation in GST. (b) Various properties are calculated to quantify the nucleation process, including the mean square displacement of the system, the number of four-membered rings, the number of four-membered rings forming planes the number of cubic-like motifs, and the number of atoms forming planes or cubes. (c) The stability of ABAB rings in GST at 600 K. Panel a and b are adapted with permission from Ref. 101. Copyright 2011 by the American Physical Society. Panel c is reproduced with permission from Ref. 108. Copyright 2012, American Association for the Advancement of Science.

crystalline nuclei are present. Under thermal fluctuation at elevated temperatures, these four-membered rings are highly mobile. Only when two rings encounter each other can they possibly form cubic motifs – the structural building unit of the crystalline counterpart. Throughout the first and second stages, the number of four-membered rings and cubic structures fluctuates around certain values, indicating that the nucleation period shows significant stochasticity, which accounts for the long incubation required for nucleation.

Note that the fast crystallisation of these small-scale models (containing ≈200 atoms) is largely affected by the finite-size effect, and in fact, nucleation might take several nanoseconds to form critical nuclei. To accelerate this process, advanced

sampling methods (e.g., metadynamics) can be used together with AIMD to decrease the computational costs. Without fixing any crystalline seed, a 460-atom structural model nucleated quickly with the help of metadynamics sampling [109]. However, to enable an unbiased crystallisation from amorphous structures, longer simulation time scales and larger simulation boxes are required.

To further investigate the origin of the stochasticity in nucleation, the lifetime of four-membered rings in the amorphous GST structure at elevated temperatures was calculated [110]. As shown in Fig. 1.13f, at high temperatures, the ABAB-type four-membered rings in GST alloys have a relatively short lifetime, making them "vulnerable" under thermal vibrations. The average lifetime of these four-membered rings in GST alloys at high temperatures is estimated to be 3 to 5 picoseconds [110]. It is also suggested by AIMD simulations that a pre-embedded crystalline seed containing 58 atoms in a 460-atom GST model disappears quickly at 600 K. As a matter of fact, even though there are a significant number of four-membered rings present in the amorphous GST, the relatively short lifetime of these four-membered rings significantly increases the stochasticity of the nucleation process, leading to a long incubation in GST (e.g., tens of hundreds of nanoseconds) and thus a slow SET process in GST-based devices.

To overcome such speed bottleneck by the intrinsic stochasticity of the nucleation process, a pre-nucleation strategy has been proposed [108]. As shown in Fig. 1.14, a constant low-voltage pulse (0.3 V), lower than that used in SET operation, is applied to GST alloys for 10 nanoseconds. Immediately after this, a crystallisation pulse is then applied to trigger the crystallisation. Based on such pre-programming scheme, GST alloys can complete crystallisation within 0.5 nanoseconds, which is an order of magnitude faster than ones without the pre-applied electric pulse (Fig. 1.14).

AIMD simulations were then performed to explain the mechanisms of the pre-programming scheme [108]. The constant low-voltage pulse induces a preliminary ordering process in the GST alloy, leading to the formation of crystalline seeds with

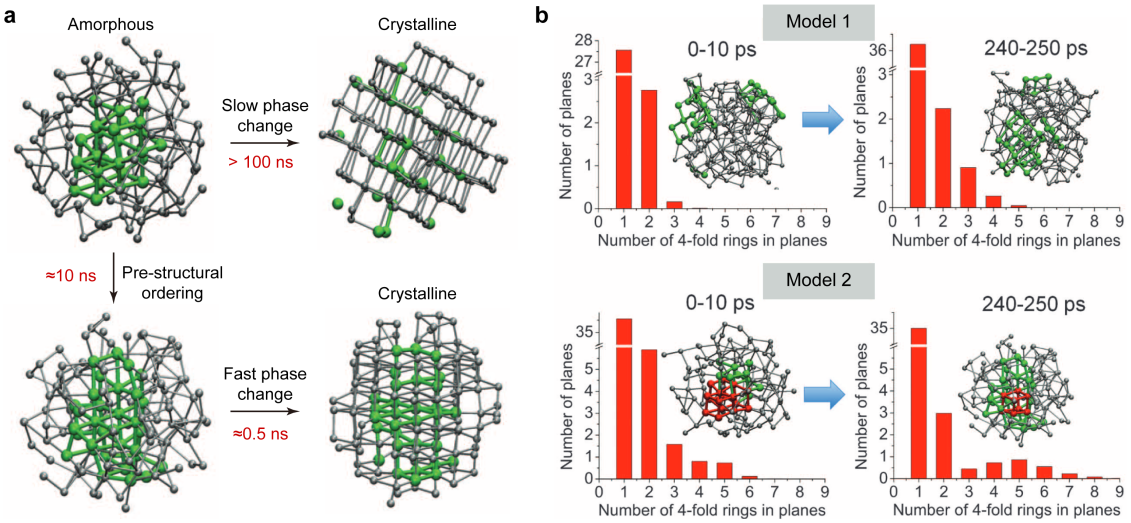


Figure 1.14: (a) A sketch of the pre-structural ordering scheme showing that a small pre-applied electric pulse of ≈10 ns can help accelerate the crystallise process. (b) Quantitative analyses of crystalline planes found in two structural models before and after the pre-programming process. An evident structural ordering process can be observed. The figure is adapted with permission from Ref. 108. Copyright 2012, American Association for the Advancement of Science.

considerable sizes in the pre-programming stage. After that, the subsequent crystallisation pulse can thus quickly crystallise GST alloys based on these pre-formed nuclei. However, note that this pre-programming strategy does not effectively increase the write speed of devices, because each write operation still requires a 10-ns pre-programming operation.

To further decrease the crystallisation time within one nanosecond, the intrinsic stochasticity should be eliminated for a faster incubation process. As discussed above, the key of incubation is the formation of four-membered rings. Hence, increasing the lifetime and chemical stability of these ABAB four-membered rings is the key. To achieve this goal, Sb_2Te_3 was used as the parent compound to avoid large structural complexity associated with Ge-centred tetrahedral motifs. Based on materials screening, scandium (Sc) was then selected as the doping element. The resultant $Sc_0.2Sb_2Te_3$ (SST) alloys showed a ultrafast crystallisation speed of only 0.7 nanoseconds, which is an order of magnitude faster than GST alloys.

AIMD simulations were carried out to understand the ultrafast crystallisation

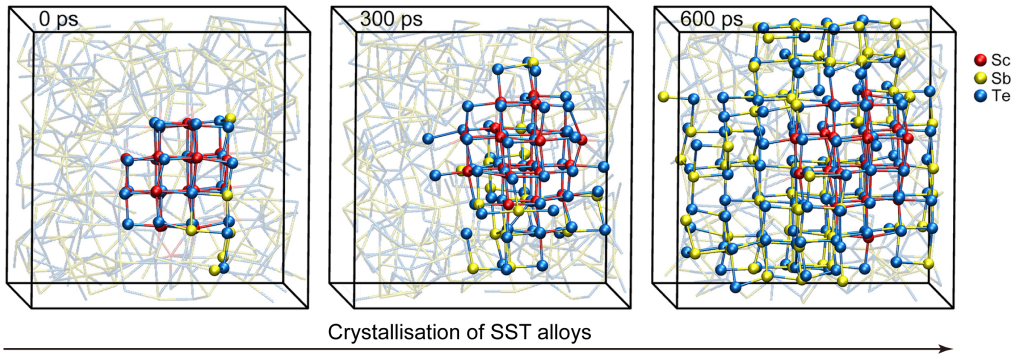


Figure 1.15: The AIMD simulation (428 atoms) of the ultrafast crystallisation process of SST alloy. A crystalline seed was pre-fixed serving as a nucleus of the subsequent crystallisation. The figure is adapted with permission from Ref. 110. Copyright 2017, American Association for the Advancement of Science.

process in SST using a simulation box of 428 atoms. As shown in Fig. 1.15, during the crystallisation at elevated temperatures, the Sc atoms in SST form strong chemical bonds with Te. The presence of chemically strong Sc-Te bonds largely stabilises the four-membered rings and cubic motifs in the SST alloy, significantly increasing their lifetime in the incubation process. Statistical analysis shows that the lifetime of Sc-including four-membered rings in SST exceeds 50 picoseconds, which is substantially longer than that of the four-membered rings in GST. Hence, the stochasticity in the nucleation process of SST is largely reduced, leading to an increased nucleation speed and thus an accelerated crystallisation process.

To better understand the role of Sc doping in the nucleation process, AIMD simulations were carried out on the parent phase of SST alloys, i.e., Sc_2Te_3 and Sb_2Te_3 [111], using ≈200 atoms. Due to large charge transfer between Sc and Te, almost no homopolar bonds were found in amorphous Sc_2Te_3 alloys, leading to the formation of complex higher-membered rings (e.g., six- and eight-membered rings) [111]. These complex rings are different from the cubic-like motifs in the recrystallised counterpart, and therefore too much Sc concentration will increase structural complexity and stochasticity, which slows down the nucleation [111].

Atomistic simulations of crystal growth

When it comes to growth-dominated PCMs, e.g., AIST alloys, AIMD simulations have also been used to explain the origin of its slow nucleation. AIMD simulations show that amorphous AIST has a broad distribution of ring sizes, ranging from three- to seven-membered rings, with five-membered rings being the most prevalent (accounting for 15% of the total). The presence of abundant five-membered rings makes the amorphous network very different from the building units of the recrystallised counterpart [112]. Hence, the structural dissimilarity between amorphous and crystalline AIST leads to the slowed-down nucleation. As a result, AIST tends to crystallise through crystal growth rather than nucleation.

To investigate the intrinsic reasons behind the rapid growth of AIST at high temperatures, Mazzarello et al. utilised AIMD simulations to study the growth process of AIST [113]. To mimic the crystal growth of AIST that starts with an existing crystalline seed, several crystalline [0001] layers were fixed during the melt-quench simulation to create a crystalline-amorphous interface. As shown in Fig. 1.16a, the structural model consisted of 810 atoms in a box of 2.2 × 2.1 × 5.1 nm^3. When annealing at ≈585 K, the growth proceeded quickly at two crystalline-amorphous interfaces, and the growth completed at the interface within a few hundred picoseconds. Notably, no significant crystalline nuclei were observed in the disordered region during the growth process, reflecting a growth-driven crystallisation behaviour.

The presence of the amorphous-crystalline boundary significantly facilitated growth, enabling a large growth rate at elevated temperatures. The growth rate can be estimated from the AIMD simulations – which was ≈7.8 m s^{-1} at 585 K, agreeing well with the experimental measurements at the same temperature [87].

To further investigate the structural and dynamic properties during the growth process, the per-atom crystallinity was quantified using the order parameter q_4^{dot}. A detailed introduction of characterising per-atom crystallinity is given in Section 3.5. The q_4^{dot} parameter ranges from 0 to 1, where a value closer to 0 indicates a

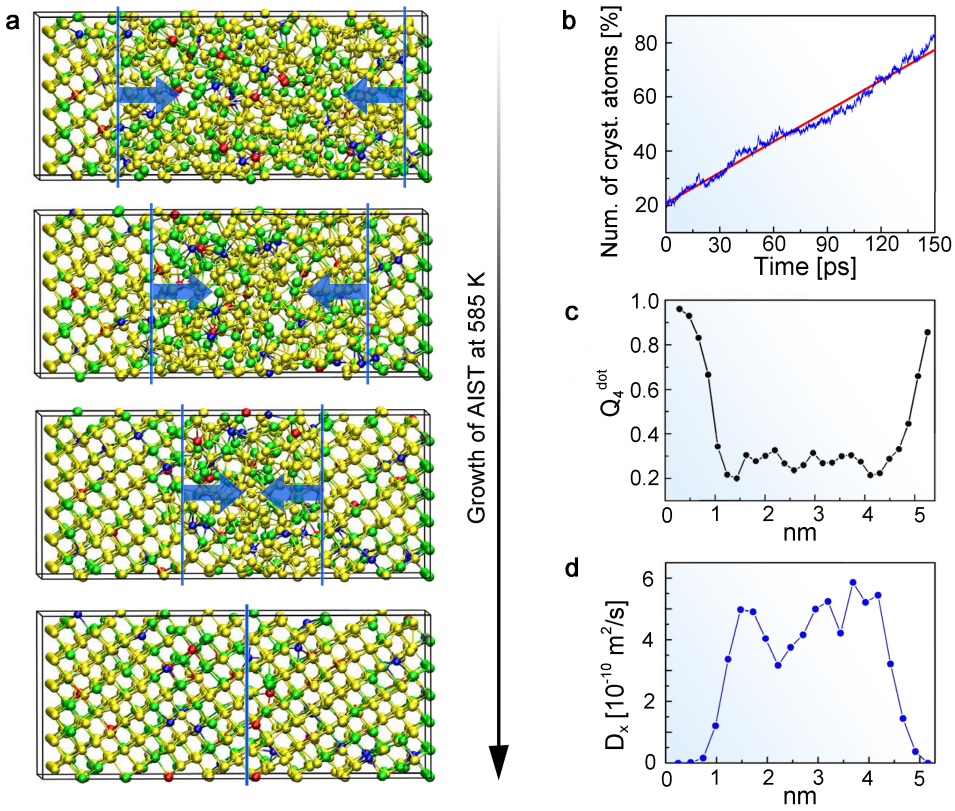


Figure 1.16: (a) AIMD simulations of crystal growth from crystalline–amorphous boundaries in $Ag_4In_3Sb_{67}Te_{26}$ (AIST) (containing 810 atoms) at 585 K. (b) Number of crystalline-like atoms formed during interface growth. (c) The profile of the bond order parameter (Q_4^{dot}) for the snapshot at the initial stage of crystallization, indicating a narrow interface of less than 1 nm. (d) The diffusion coefficients along the growth direction (D_x) show a sub-nanometre narrow interface, corresponding to a rapid reduction in atomic mobility from the amorphous region to the crystalline front. The figure is adapted with permission from Ref. 113. Copyright 2019, the Authors. Published by Springer Nature.

more disordered atomic environment that might be found in the amorphous network whereas a value closer to 1 indicates a more ordered environment, approaching a perfect octahedral motif.

Figure 1.16c shows the distribution of the calculated Q_4^{dot} along the [0001] direction at the initial stage of the growth process, which was computed by the averaged q_4^{dot} values within a given slab. The amorphous and crystalline regions can be clearly separated by Q_4^{dot} parameters. A narrow transition region from the ordered crystalline region to the disordered region is found, in which the average Q_4^{dot} value of the atomic layers decreases rapidly. The narrow growth interface has a width of less

than 0.8 nm, which is largely different from the growth behaviour observed in some metals [113].

In addition, the diffusion coefficient of atoms in different slabs of the structural model along the [0001] direction was computed via the mean square displacement (MSD) (Fig. 1.16d). It is found that the diffusion coefficient of atoms in the amorphous matrix was 5.0×10^{-10} m^2 s^-1 at 585 K. The high diffusion rate indicates high atomic mobility, which is key to a fast growth process. However, as atoms approached the crystalline-amorphous interface, their diffusion coefficients decreased significantly, which suggested a high probability for free atoms to attach to the crystalline matrix within the narrow growth interface.

Hence, the growth process of PCMs exhibits three important characteristics: high atomic mobility, a narrow growth interface, and an efficient atomic attachment process. Specifically, the high atomic mobility serves as the kinetic basis for rapid growth. The narrow interface facilitates frequent and rapid collisions of atoms with the crystalline-amorphous interface. The high probability of atomic attachment ensures that atoms from the disordered structure effectively adhere to the growth interface, enabling rapid and effective growth.

Note that the fast growth process was also observed in the AIMD simulations of the crystallisation in GST alloys [109, 114]. Starting from a similar templated structural model, i.e., a melt-quenched $Ge_2Sb_2Te_5$ model (containing 540 atoms) with a prefixed crystalline matrix serving as the growth template, the growth proceeded quickly at the crystalline-amorphous interface [109]. More importantly, three important characteristics as mentioned above were also observed in the growth of GST alloys. The estimated growth rate was $\approx$1 m s^{-1} which agreed reasonably with the experimental data. More interestingly, a large-scale structural model, which contains 1,008 atoms, was used to simulate the growth process of $Ge_1Sb_2Te_4$. This size has already approached the limit of AIMD simulations. A detailed discussion of this large-scale growth simulation is given in Section 3.5.

Atomistic simulations of iterative RESET and cumulative SET

AIMD simulations have also been carried out to simulate two key modes in in-memory computing, i.e., iterative RESET and cumulative SET. The key to these operations is the implementation of a gradual amorphisation or crystallisation using a series of RESET or SET pulses (Fig. 1.11). To model the pulsing scheme that implemented spike-timing-dependent plasticity with PCM cells [31], different pulse widths and amplitudes were considered in AIMD simulations [115]: different pulse amplitudes corresponded to different anneal temperatures (Fig. 1.17a) whereas different pulse widths were represented by different annealing time in AIMD simulations (Fig. 1.17b).

The cumulative SET process was nicely described in AIMD simulations (Fig. 1.17). A series of pulses with a constant pulse width (corresponding to the same annealing time) but different pulse amplitudes (corresponding to different annealing temperatures) were added on the structural model in AIMD (Fig. 1.17a). The low-temperature annealing of the melt-quenched amorphous model led to the formation of small but stable crystalline clusters (Fig. 1.17c). As the pulse amplitude increased, a subsequent growth was observed starting from these stable crystalline clusters. Moreover, when a series of pulses with a constant amplitude (corresponding to the same annealing temperature) but different pulse widths (corresponding to different annealing durations) were added in AIMD, a similar gradual crystallisation was also observed, which was captured by the number of crystalline rings (Fig. 1.17a and b). However, the very short annealing time had very few effects on the structural ordering process in the cumulative SET.

Although the cumulative SET process was reproduced in AIMD simulation by adjusting the annealing temperature and time (corresponding to the pulse amplitudes and widths, respectively), the iterative RESET process, i.e., a gradual amorphisation process, can hardly be simulated in the structural model of only several hundred atoms, due to the finite-size effect. A large RESET pulse, i.e., a high-

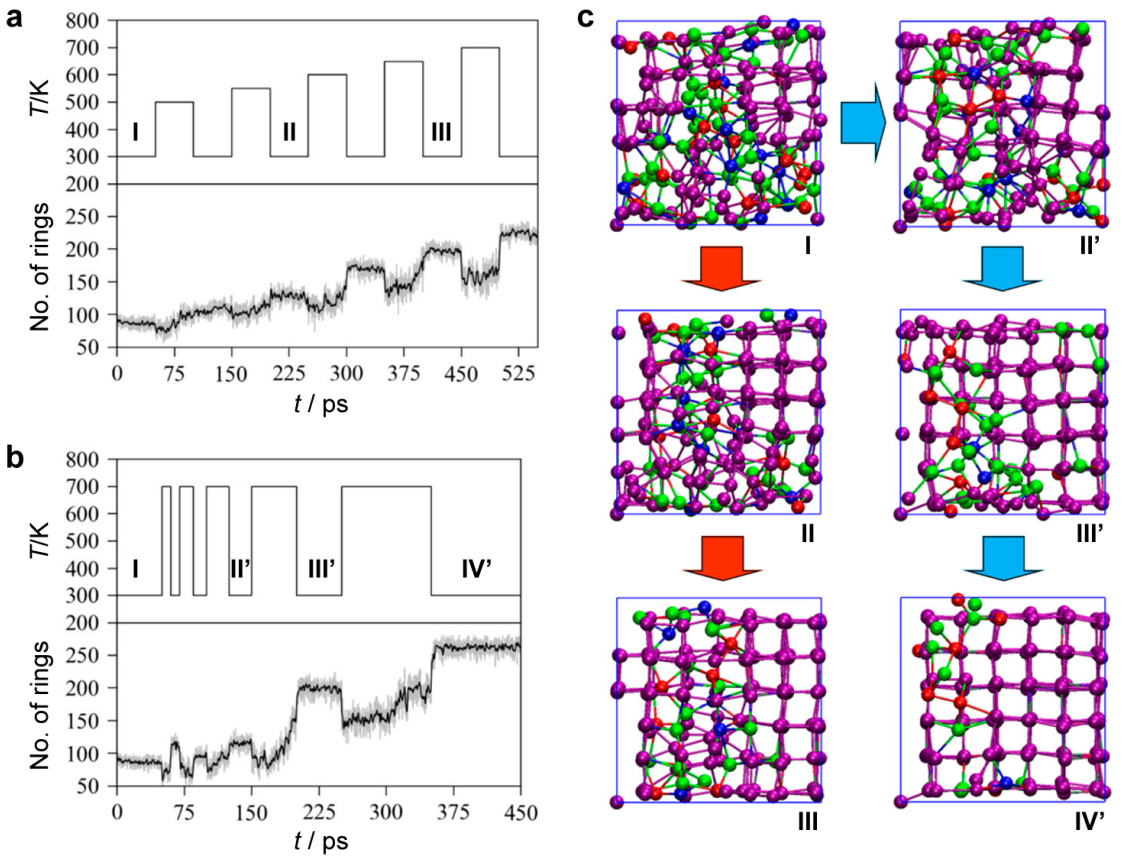


Figure 1.17: Simulations of cumulative SET process. Two different schemes are shown including (a) a series of SET pulses with a constant width but different amplitudes, and (b) a series of SET pulses with a constant amplitude but different widths. The number of rings indicates a gradual crystallisation process under a series of SET pulses. Six configurations during the cumulative SET process are shown in (c), marked as **I**, **II**, **III**, **II'**, **III'**, and **IV'**. Atoms are rendered in different colours: Ge, blue; Sb, red; Te, green. The figure is adapted with permission from Ref. 115. Copyright 2015, American Chemical Society.

temperature heating process, will quickly melt the whole structural model and all atoms become disordered in a collective way.

Note that the annealing time for each operation pulse in the AIMD simulations ranges from several to tens of picoseconds, which is much shorter than the nanosecond pulses employed in the realistic PCM-based devices [116]. Hence, large-scale (e.g., real device-scale) MD simulations over nanoseconds are required to better model the in-memory computing applications.

Moreover, as the operating pulses are modelled as constant annealing processes, it is impossible to simulate the overlap of separated pulses in neuromorphic computing

[31], failing to describe more complex in-memory computing schemes. Hence, the discrepancies between the temperatures used in the simulations and the effective temperatures induced by Joule heating from the low-voltage pulses applied in PCM-based neuromorphic devices make it even harder to fully simulate the realistic in-memory computing operations using AIMD.

1.5 ML-driven atomistic simulations of phase-change materials

Despite the important contributions made by AIMD simulations for PCM research as mentioned above, it is not feasible to carry out large-scale simulations (e.g., millions of atoms) over nanoseconds (corresponding to millions of MD steps) and with sufficient statistical sampling to study different physical properties and complex programming processes, given the heavy computational loads. For example, an AIMD simulation of more than 400 atoms over 1 ns would already take more than half a year real-world time using a high-performance supercomputer [105, 109].

Classical force fields have largely increased efficiency by several orders of magnitude as compared to the AIMD simulations, which avoid expensive electronic structure calculations and can be used for large-scale atomistic modelling. However, the complex interactions between different atomic environments in PCM systems make the classical potentials very challenging to accurately simulate. More importantly, when three or more elements are considered in the same system, the empirical potentials usually fail in describing such chemical complexity. A Tersoff potential was developed for GeTe [117], which can provide an acceptable description of the crystalline, liquid, and amorphous phases of GeTe. However, when it comes to more complex phase transitions with crystalline-amorphous interfaces, the Tersoff potential might become less feasible.

An important theoretical advance in the past decade is the development and use of machine-learning (ML) interatomic potentials, which enable comprehensive

studies of different properties of a wide range of PCMs [118]. To achieve both accuracy and efficiency, ML potentials provide a promising solution to the conundrum with quantum-mechanical accuracy and an efficiency almost comparable to classical potentials.

ML potentials are a mathematical representation of $3N$-dimensional potential energy surface describing the relationship among the spatial coordinates of N atoms, total energy, interatomic force and other properties [118]. In contrast to classical potentials, ML potentials do not make prior assumptions about the exact mathematical form of interatomic interactions in the systems, but instead are "learned" from a large reference database. A detailed introduction of ML potentials is given in Section 3.1 of this thesis, and in the following some recently developed ML potentials for PCMs are introduced.

1.5.1 Neural-network-based potentials for GeTe

The first ML potential for PCMs was developed for GeTe using the neural-network framework (Fig. 1.18a) [119]. Around 30,000 configurations of 64, 96, and 216 atoms were included in the training database. To construct such a large reference database, an initial training database of only 5,000 structures was first obtained, which only contained the crystalline, liquid, and melt-quenched amorphous structures of GeTe under different external pressures, as well as the crystallisation snapshots from metadynamics simulations [119]. The initial database was then gradually extended by adding random structures of local distortions and slightly varied stoichiometries. A mathematical representation, known as atom-centred symmetry functions (ASCF) [120] was then used to incorporate the structural information into a neural network (NN) algorithm. The resultant NN potential can accurately describe the different complex phases of GeTe [119]. Moreover, the NN potential has also been used to study the crystallization kinetics [121–123] and thermal conductivity [124, 125] of GeTe using large-scale structural models (containing thousands of atoms).

As shown in Fig. 1.18b, a 4,096-atom structural model of GeTe was used for

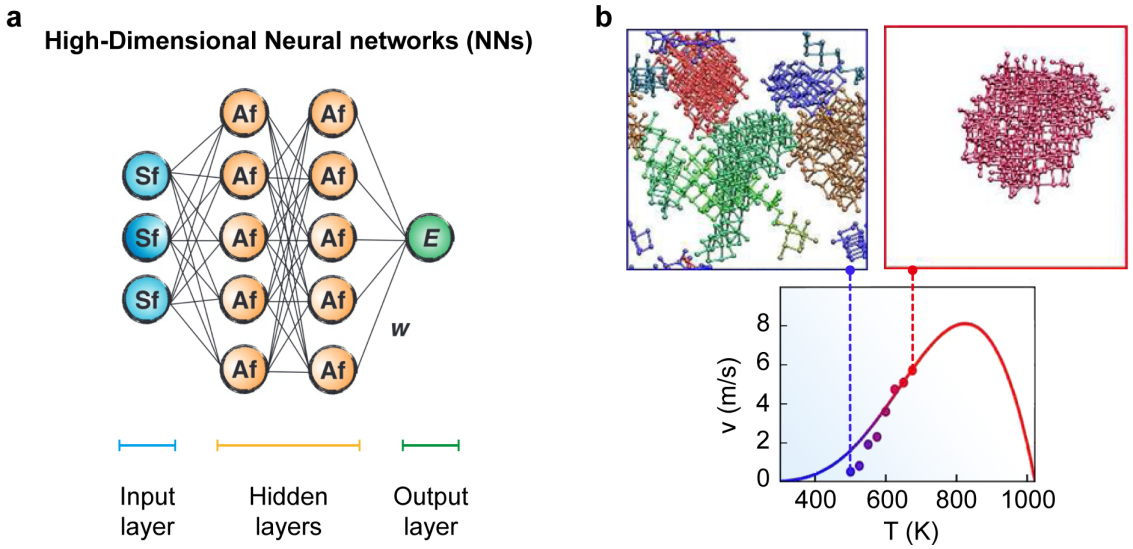


Figure 1.18: (a) Neural network (NN) can be used to construct an ML potential starting from a dataset of DFT-computed energies of training configurations. "Sf" represents the atom-centred symmetry functions (ASCF) as the input of NN. "Af" indicates the activation functions used in the fitting. "E" is the energy of an individual structure. The weights, w, of the NN potential are updated by back-propagation. (b) The calculated growth speed of GeTe from 500 K to 675 K. Single and multiple crystalline seeds are found at 675 K and 500 K, respectively. Panel a is adapted with permission from Ref. 126. Copyright 2019, the Authors. Published by Springer Nature. Panel b is adapted with permission from Ref. 121. Copyright 2013, American Chemical Society.

the study of crystallisation using the NN potential. This size is much larger than the limits of AIMD capacity and reduces the finite-site effects. In ML-driven MD (MLMD) simulations, nucleation and subsequent crystal growth were observed at elevated temperatures [121]. Growth rates were determined to be on the order of a few metres per second, which are similar to those found in GST [121]. In addition, at lower temperatures (e.g., 500 K) close to the glass transition temperature, a competition between nucleation and growth was observed, whereas at higher temperatures (e.g., 700 K), the crystal growth was more dominant [122].

More importantly, the NN potential can be further extended to study other complex scenarios, e.g., GeTe nanowires. To obtain the extended NN potentials, new structural models that only contain several hundred atoms and can represent the target systems are collected. These new structures were then added back to the original reference database for the training of new NN potentials [127]. As shown

in Fig. 1.19a, a partly amorphised GeTe nanowire model (9 nm in diameter) with 16,540 atoms was simulated using an updated version of the original NN potential for GeTe [127]. The amorphous part of the nanowire model shows a lower fraction of Ge-Ge chains than that in bulk amorphous GeTe, which could partly explain the reduced drift coefficient in nanowires [128]. Moreover, the recrystallisation of this partially melted GeTe nanowires was estimated to be slower than that of the bulk GeTe (Fig. 1.19b), and the melting point of the crystalline nanowire model is about 100 K lower than that of bulk GeTe [127].

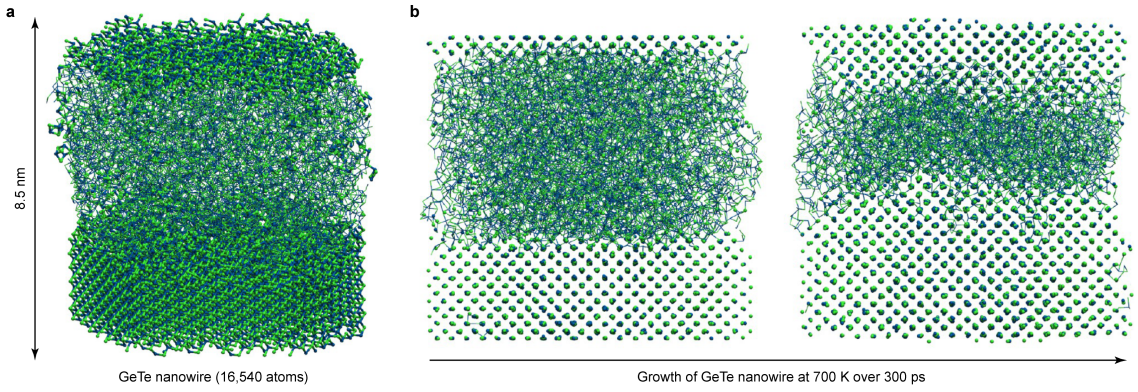


Figure 1.19: NN-based large-scale MD simulations. (a) The atomistic modelling of a GeTe nanowire with a partly amorphised fragment. The model contains 16,540 atoms with a diameter of ≈9 nm and a nanowire length of ≈8.5 nm. (b) Recrystallisation of a partially melted GeTe nanowire at 700 K over 300 ps. The figure is adapted with permission from Ref. 127. Copyright 2017, American Chemical Society.

In addition, by adding interface structures of the GeTe-$TiTe_2$ superlattice, the resultant NN potential for GeTe-$TiTe_2$ heterostructure was used to study the crystallisation of a nanoconfined superlattice [129]. An increased crystallisation temperature was found in the simulations of GeTe-$TiTe_2$ heterostructure as compared to the experimentally identified crystallisation of the Sb_2Te_3-$TiTe_2$ heterostructure which was originally proposed to be a good material candidate for in-memory computing [47].

Note that ML potentials have larger transferability (or extrapolation ability) to be used in other complex scenarios that have not been considered in the training of potentials, as compared to empirical potentials. Hence, the diverse training struc-

tures of the NN potential for GeTe have enabled a transferable ML model which can be used as an off-the-shelf theoretical tool in other studies of GeTe, including the pressure-induced reversal of Peierls-like distortions in amorphous GeTe [130], and the configurational entropy and the fragility of supercooled liquid GeTe [131].

Notably, NN framework has also been used in other PCM systems. An NN potential was developed for elemental antimony (Sb) [132], which investigated the hindered crystallisation in a nanoconfined Sb thin film (3–5 nm thick) in stark contrast to the ultrafast crystallisation in bulk Sb [133]. It is found that the largely reduced crystallisation speed resulted from the effect of nanoconfinement on the fast β relaxation while the slow α relaxation is essentially unaffected [133]. In addition, an NN potential of Ge-rich Ge_xTe alloys has been developed which covered a wide range of binary Ge–Te compositions [134]. This NN potential accurately described the structural properties of crystalline, amorphous, and liquid Ge-rich Ge_xTe alloys, as well as the crystallisation process at elevated temperatures. MD simulations of 19,200 and 30,000 atoms over up to 40 ns were carried out [134], showing that phase separation occurs at a relatively short time scale of a few ns in Ge_2Te at high temperature (e.g., 600 K).

1.5.2 ML potentials for Ge–Sb–Te alloys

Due to more complex chemical environments in the ternary Ge–Sb–Te alloys, developing ML potentials for GST is more challenging. In 2018, the first ML potential of $Ge_2Sb_2Te_5$ was developed using the Gaussian approximation potential (GAP) framework [135], a kernel based regression method that will be introduced in more details in Section 3.5. To encode the local structural environments, the smooth overlap of atomic position (SOAP) was used. This GAP potential included only about 3,300 training structures, including elemental, binary, and ternary crystals, as well as melt-quenched liquid amorphous configurations of $Ge_2Sb_2Te_5$. Comprehensive structural validations were carried out, showing that the GAP potential can accurately describe the complex structures of $Ge_2Sb_2Te_5$. Large-scale MD simulations

were carried out with a model size of up to 24,300 atoms [136], much larger than the length scale of conventional AIMD simulations (Fig. 1.20a). A crystallisation of amorphous $Ge_2Sb_2Te_5$ was simulated using the GAP model (Fig. 1.20b). Moreover, due to largely increased efficiency of the GAP potential, sufficient statistical sampling over 30 independent melt-quench MLMD runs was obtained to investigate the mid-gap states in amorphous GST (Fig. 1.20c and d) [137], which served as a promising protocol to elucidate the underlying aging paths in amorphous GST [138].

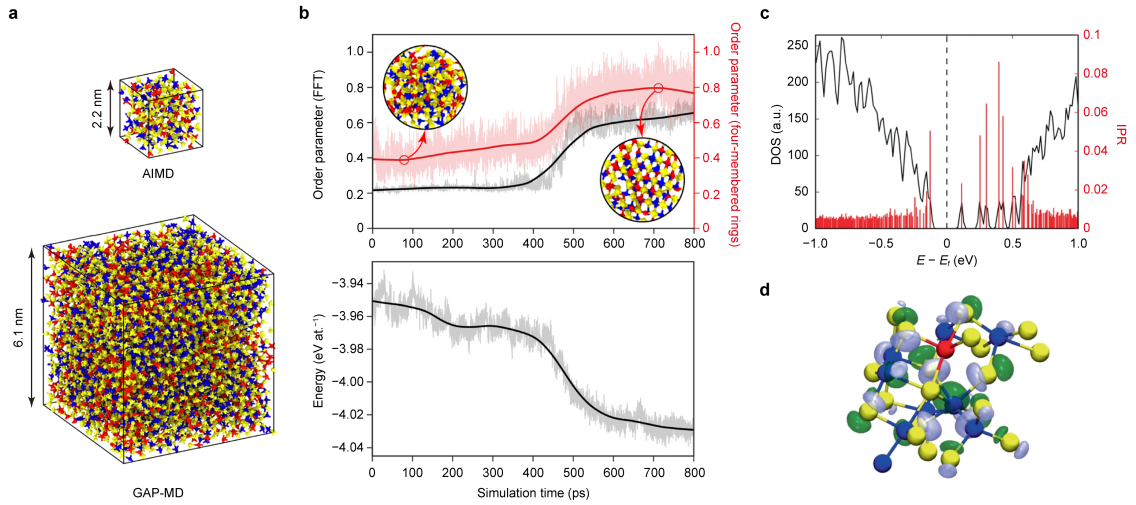


Figure 1.20: GAP potential for $Ge_2Sb_2Te_5$. (a) A comparison of a small AIMD-based amorphous GST model (containing 315 atoms) and a large GAP-based amorphous GST model (containing 7,200 atoms). (b) The crystallization of amorphous $Ge_2Sb_2Te_5$. The upper panel shows increased structural order, quantified using the second largest component of the three-dimensional Fourier transform of the atomic density and the ratio of four-membered rings. Representative structural fragments are shown and illustrate the transition from an amorphous to a partially crystallised structure. The lower panel shows the potential energy of the system. (c) Total electronic density of states (DOS) and the inverse participation ratio (IPR) values for the Kohn–Sham orbitals of an amorphous $Ge_2Sb_2Te_5$ model containing 900 atoms. (d) The atomic structure and molecular orbital associated with the mid-gap electronic state found in the band gap of the 900-atom glass model. Panel a and b are adapted with permission from Ref. 135. Copyright 2018, American Chemical Society. Panel c and d are adapted with permission from Ref. 137. Original figures published under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

Moreover, based on the GAP potential, large-scale MLMD simulations have been carried out to study the partial melting behaviour of $(GeTe)_m$-$(Sb_2Te_3)_n$ superlattice [139]. A sparse-nucleation-and-growth governed melting behaviour was found, which was triggered by the migration of atoms into the van der Waals (vdW)-like gaps of

the superlattice. The hopping of atoms created local structural disorder at elevated temperatures and thus resulted in the starting of the subsequent melting across the liquid-crystalline interfaces in the superlattice. This melting behaviour was similar to that in hexagonal GST based on the previous AIMD results [95], but was largely different from that observed in the bulk rock-salt GST structure, in which the randomly distributed atomic vacancies gathered and formed larger voids [139].

More recently, a new NN potential has been developed for $Ge_2Sb_2Te_5$ [140]. However, due to the largely increased chemical and structural complexity of $Ge_2Sb_2Te_5$ in comparison to GeTe, around 180,000 training structures were collected for this NN potential. Numeric and structural validations were also carried out, showing good agreement with the DFT and AIMD references [140]. To study the crystallisation of $Ge_2Sb_2Te_5$, a structural model containing 12,960 atoms was constructed. Due to the alleviated finite-size effects, the stochastic nature of the crystallisation was captured in the large-scale MLMD simulations. Multiple nuclei were found at lower temperatures (e.g., 500–650 K), and the subsequent fast growth led to polycrystal structures upon full recrystallisation. Moreover, a high-temperature annealing simulation at 680 K was carried out using a 27,000-atom polycrystal model over 22 ns, showing a coarsening of crystalline grains [140].

1.5.3 Graph neural-network models for Sb–Te alloys

In addition to the NN and GAP potentials which are largely dependent on the choice of local structural descriptors (i.e., ASCF in NN and SOAP in GAP), another newly developed framework for ML potentials is the implementation of (message-passing) graph neural-network algorithm [141]. Taking atoms as nodes and chemical bonds as edges between nodes, a three-dimensional atomic structure can naturally be transferred into a graph, leading to graph-based structural descriptors.

Recently, Wang et al. developed a graph convolutional neural network (GCNN) for six binary Sb_xTe_y alloys, covering a wide range of compositions (ranging from 7:1 to 1:2) [142]. Comprehensive numeric and structural validations have been carried

out, and the GCNN potential showed small training errors and good agreement in the structural properties of different phases of Sb–Te alloys as compared to the AIMD references. The accurate description of different Sb–Te compositions suggested great chemical transferability of graph-based ML framework [142].

Note that the graph-based ML potentials have been shown to exhibit great data-efficiency (that is, they require fewer training data points as compared to other non-graph ML potentials) and excellent extrapolation ability, which serve as a promising framework for general foundation models [143, 144]. Hence, to develop potentials for complex PCM systems, including multi-component systems with more than four elements (e.g., GST-$TiTe_2$ heterostructure [47]), complicated interfaces or surfaces (e.g., nanowires, nanoparticles), and other geometrical forms with strong confinements (e.g., thin Sb films encapsulated ZnS and SiO_2 [145]), the graph-based ML framework will be of great use.

1.6 Challenges for atomic-scale modelling of PCM-based devices

Atomistic modelling has played an important role in understanding fundamental properties of PCMs and can be potentially used in the simulations of real-world PCM-based devices in the future. In Section 1.4.1, AIMD simulations have been shown to provide useful insights into complex liquid and amorphous structures of different PCM systems as well as complicated phase transition processes (melting and crystallisation). However, AIMD simulations are too expensive, and running large-scale nanosecond-scale AIMD simulations might take several months of real-world time. Such computational costs make AIMD unable to access both the length and time scales to simulate the programming processes in PCM-based devices.

As mentioned in Section 1.5, the field of ML potentials is under very active development in the past decade. MLMD provides quantum-mechanical accuracy, and more importantly, dramatically increased computational efficiency (which is several

orders of magnitude faster than AIMD), allowing atomic-scale insights into fundamental properties in large-scale structural models or long simulation time scales (up to nanoseconds). Such advances in accuracy and speed have made MLMD a suitable theoretical tool to model the programming of PCM-based devices at the atomic scale. However, some challenges remain to fully enable atomistic simulations of PCM-based devices. Several primary challenges still exist and will be discussed in the following.

1.6.1 Exploring a diverse compositional space

A primary difficulty of PCM-based device simulations is that ML potentials are not sufficiently transferable to a wide range of chemically and structurally different atomic environments. The trained potential can only be used for specific chemical compositions. Despite that the GAP potential originally developed for $Ge_2Sb_2Te_5$ [135] can also be used in the simulations of Sb_2Te_3 [146] and the GST superlattice [139], it is still challenging for this stoichiometrically-specific GAP potential to cover the whole GeTe-Sb_2Te_3 pseudo binary line (i.e., the flagship GST compositions for widely used GST-based devices) and consider large compositional variation (e.g., large elemental migration during extensive cycling) in real GST devices. In particular, different GST alloys show slightly different melting and crystallisation behaviours. For example, Sb_2Te_3 tends to have a faster nucleation than GeTe, due to the different amorphous structures between Ge-rich and Ge-poor environments, thus obtaining a lower crystallisation temperature [147].

Moreover, if high pressure or surface/interface effects need to be considered for the specific composition, additional fitting and learning are necessary. For example, in real GST devices, the phase-change layers are surrounded by dielectric materials. Hence, to study the interface effects between the GST layers and the dielectric layers, additional training data are also required.

However, it is not easy to efficiently collect training data for a wide range of compositions. A large database might lead to an expensive fitting process and a risk

of heavy inference process (i.e., a decreased computational efficiency). Hence, a well-designed training protocol is required, covering multiple compositions at different temperature and pressure points, which helps to build a diverse training database.

1.6.2 Simulating programming of device

As mentioned in Section 1.3, electric or laser pulses are used to trigger the phase transition processes of PCMs. The external pulses will result in a non-isothermal heating process where an evident temperature gradient will be observed. However, in most AIMD and MLMD simulations (i.e., constant temperatures), only the isothermal heating conditions were used, due to the large complexity of implementing temperature gradients. Hence, to mimic more realistic programming conditions in PCM-based devices, the complex heating conditions need to be considered.

In fact, non-equilibrium molecular dynamics (NEMD) simulations under non-isothermal conditions were indeed carried out to compute the thermal transport properties in GeTe nanowires using MLMD [125]. Thermal conductivity can thus be computed by calculating the heat flux in the material domains. In these NEMD simulations, a hot and a cold source were defined, which were used to exchange the kinetic energies in the system. This thus suggests a possible way to simulate the additional heats, introduced by external pulses, by manually adding per-atom kinetic energy to the atoms in the programming domain.

In addition, complex multi-field coupling exists in PCM-based devices. For example, in GST-based electronic memory devices, the Joule heat is generated by electrodes, and hence, a good description of the relationship between electric field and the induced thermal filed should be provided. In FEM simulations, such relationship can be interpreted by solving the classic continuum equations. How to implement these physical relationships in ML-based atomistic modelling remains an unsolved question. Another example is the electromigration of cations and anions under electric filed, which requires an accurate description of the relationship between the external electric field and atomic charges.

More importantly, a long simulation time is needed to simulate the programming processes of PCM-based devices, because each operation (including SET and RESET) requires nanosecond-scale pulses. Given the device-scale atomistic model (tens of hundreds of nanometres in length), a large-scale MD simulation over nanoseconds is needed. Hence, an accurate and very efficient ML potential should be considered.

1.6.3 Complex device architectures

As mentioned in Section 1.2.4 and 1.2.5, two important device architectures, including the confined cells and mushroom-type devices, are widely used in PCM-based memory and in-memory computing techniques. Hence, to simulate the realistic programming environments of PCM-based devices, these two architectures should be considered. However, in most AIMD and MLMD simulations, relatively simple and periodic simulation boxes were constructed, which represented the bulk materials but largely deviated from the geometries of real PCM-based devices. Hence, a more realistic structural model should be considered in the simulations of PCM-based devices.

More importantly, complex interface interactions exist in real PCM-based devices. Different functional layers are connected with each other. To enable real device simulations, those complex interfaces need to be properly simulated.

1.7 Going beyond atomic-scale simulations

In addition to atomic-scale modelling of PCMs, other methods have also been used to model the PCM-based devices, as shown in Fig. 1.21. According to different computational strategies, the hierarchy of different computational methods can be categorised into the DFT-based methods, the analytical models, the finite-element method (FEM) models, and the compact models. The DFT-based methods are largely limited within nanometre-scale containing hundreds of atoms.

As already discussed in Section 1.4, the DFT-based simulations predict the electronic structures (e.g., energy band) and physicochemical properties of PCMs, thus

allowing for the understanding of complex switching mechanisms of PCMs and the screening of novel PCMs with optimised performance [47, 110]. In addition, AIMD simulations provide atomic-scale insights into the switching processes of PCMs, revealing the intrinsic mechanisms of complex programming.

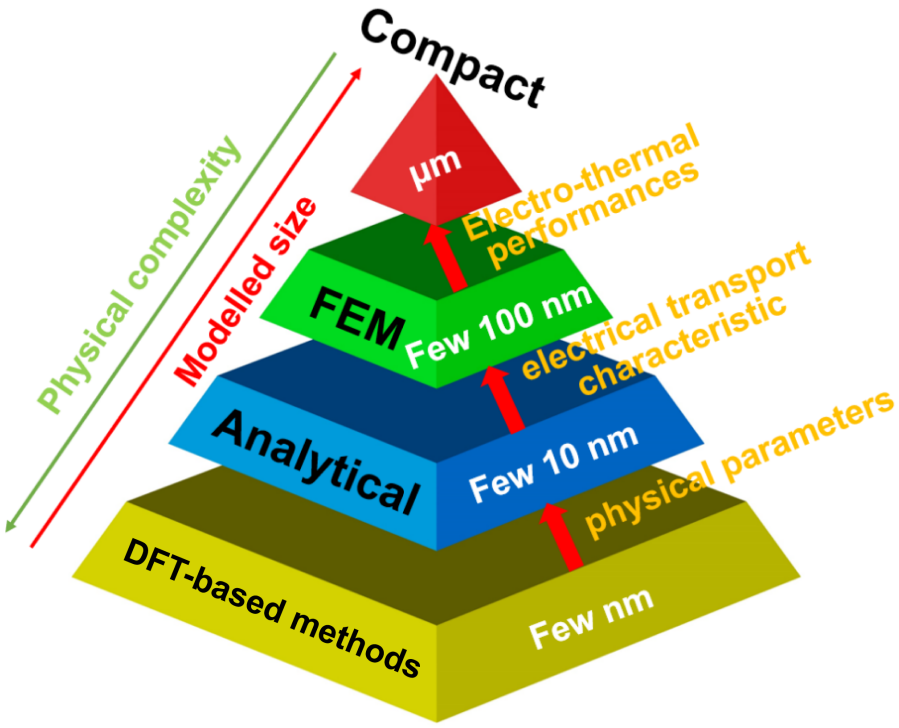


Figure 1.21: The hierarchy of different simulational methods for PCMs and PCM-based devices, based on the physical complexity and simulation sizes of the methods. Communications across the hierarchy can pass the required parameters computed from the previous level to the next level (as indicated by the red arrows). The figure is adapted with permission from Ref. 148. Copyright 2022, IEEE.

Analytical models (in a length of tens of nanometres) are used to compute the current-voltage (I-V) curves based on the electrical conduction characteristics of PCMs, enabling the simulations of the electrical transport characteristics in PCM-based devices subjected to external programming pulses. Notably, many analytical methods can adopt the computed parameters (e.g., atomic charges) from the DFT-based calculations. For example, the QuantumATK platform has provided a series of DFT-based methods [149], including pseudopotential DFT calculations using LCAO basis and plane-wave basis, as well as semi-empirical tight-binding methods [150,

151], to cover a wide range of system sizes up to thousands of atoms. Based on the DFT-computed properties, the transport engine of QuantumATK can simulate the transport properties [149].

Finite-element modelling (FEM) simulations can deal with the system size of hundreds of nanometres, approaching the realistic sizes of a single bit of PCM-based devices. As a necessary step to reduce financial and temporal costs in developing PCM-based devices, technology computer aided design (TCAD) uses numerical modelling to simulate physical properties and optimize performances of devices [148, 152]. In particular, FEM simulations are one of the key components of TCAD. By setting up a two- or three-dimensional mesh of real-world device geometry, FEM considers different device components and complex electrothermal responses, and numerically solves differential equations according to physical laws [148]. Different results can be obtained from FEM simulations, including the temperature distribution and the mechanical stress during a RESET operation induced by the density variation upon phase change (Fig. 1.22) [153]. In particular, the computed temperature distribution of the programming cells allows for accurate control of heat density [154] and assesses the robustness of devices against thermal disturbance during nanosecond operations [155].

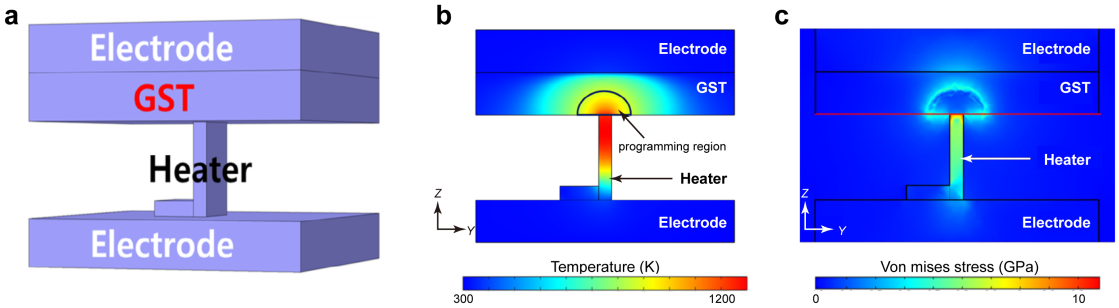


Figure 1.22: FEM simulations for PCM-based devices. (a) The three-dimensional device geometry. (b) Temperature distributions of the cross-section of the model after a RESET pulse. (c) The stress distributions of the cross-sections after a RESET pulse. The figure is adapted with permission from Ref. 153. Copyright 2020 Wiley-VCH.

However, built on continuum mechanics, FEM simulations lack atomic-level details of complex phase-change processes, and the resulting simulational errors might accumulate throughout the modelling of extensive switching. For example, traditional FEM simulations identify the amorphisation process only based on the experimental melting point [156]; however, due to overheating effects, subnanosecond annealing above the melting point might not immediately result in a melted region in RESET operations [157]. More importantly, crystallisation of PCMs exhibits prominent stochasticity [101, 105]. Despite the fact that comprehensive mathematical models have been applied in simulations of the nucleation and growth process [156, 158], FEM simulations cannot reflect the intrinsic crystallisation stochasticity of PCMs, failing to accurately describe grain morphology during SET operations. Note that the communications across the hierarchy could be very useful. Hence, given the requirements of atomic-level information in PCMs, FEM simulations might be able to combined with atomistic modelling for real device simulations. Considering the computational efficiency, combined MLMD and FEM modelling might be a good option for such applications.

Compact models, on the top of the simulation hierarchy, can be used to simulate the entire memory devices (in a size of up to micrometers), including all functional units and circuits, and provide circuit response when applied external bias (e.g., the I-V response of the whole circuit). In fact, compact models in essence belong to analytical models, but cover a larger length scale, which simplify the whole memory device into equivalent circuits and imitate the phase-change processes in an analytical way. Note that most of the analytical models were developed based on the "threshold switching" behaviour (as discussed in Section 1.2.3). The electronics-like model is one of the simplest compact models for PCMs. In this model, a single PCM-based memory unit will be simplified into a functional unit on a circuitry scale. Only the static resistance, voltage, and current of different states of PCMs are considered, neglecting the dynamical effects, e.g., the temperature-dependent electrothermal ef-

fects and the phase transition processes. A modification can be performed on the top of the electronics-like model, via a series of physically-informed equations, leading to the physics-like model, which includes the dynamical phenomena during the operations of PCM-based devices [148]. However, similar with the problems in FEM simulations, all compact models lack atomic-level information of phase transitions, thus failing to capture the stochastic nature of PCMs.

Across the whole hierarchy of simulational methods for PCMs and PCM-based devices, the FEM simulations and DFT-based simulations are two most important components to understand fundamental properties of PCMs and the microscopic programming mechanisms of PCM-based devices. However, unlike the study of conventional silicon-based devices (e.g., SRAM, DRAM, and SSD), the stochastic nature of PCMs during the phase change processes is of great importance to be captured in modelling. Hence, this thesis will mostly focus on the discussion of atomistic simulations of PCMs and PCM-based devices.

Chapter 2

Thesis outline

The aim of this thesis is to showcase the usefulness of machine-learned interatomic potentials in the study of complex solids, to provide useful protocols to develop ML potentials, and to study fundamental properties and switching mechanisms of different phase-change materials. A particular focus of this thesis is to demonstrate the device-scale simulations of Te-based selector and GST-based memory devices using ML potentials, which paves the way for full device simulations in the future.

Chapter 3 provides an overview of the different methodologies used in this thesis, focusing on the development of machine-learned interatomic potentials. Two key components for a good ML potential will be discussed, including the generation of a reference database and the optimisation of hyperparameters of ML models. Two ML potential frameworks will be highlighted, including the Gaussian approximation potential (GAP) and the atomic cluster expansion (ACE), both of which will be largely used and discussed in the following chapters. Next, I will showcase the investigation of the structural properties of amorphous red phosphorus using a GAP potential, which aims to demonstrate the usefulness of ML potentials in the study of structurally complex inorganic solids. Moreover, to quantitatively study the crystallisation and melting processes of PCMs in the following chapters, computational methods of quantifying per-atom crystallinity will be discussed.

Chapter 4 focuses on the study of tellurium (Te) – an elemental system with large structural complexity. A two-step iterative training protocol will be introduced, including standard iterations and, more importantly, domain-specific iterations. I will show how domain knowledge can be used in the training process of ML potentials, which helps the potentials to learn about the key structural information regarding phase transition processes. Given the lack of in-situ atomic-scale experimental characterisation of crystallisation in Te at nanosecond timescale, MLMD simulations will

be used to help understand the intrinsic mechanism of crystallisation and chirality transfer of Te solids. In addition, the large-scale MLMD simulations over tens of nanoseconds, comparable to the real device size, shed light on the full-loop programming process of Te-based selector devices and explain their working mechanisms.

Chapter 5 shows the study of ternary Ge–Sb–Te alloys, both chemically and structurally more complex than P and Te. I will show how to use the two-step iterative training protocol to collect the training structures, leading to a chemically and structurally diverse reference database. Comprehensive validations on the potential models will be carried out, showing both numeric accuracy and structural agreement as compared to the DFT and AIMD reference. More importantly, given complex programming conditions in GST-based devices, I will discuss how to model complex switching scenarios using MLMD, and in particular, I will showcase how to implement the non-isothermal heating conditions in MLMD simulations, which is key for real device operations. A device-scale structural model of GST containing more than half a million atoms will be shown, which demonstrates the possibility and usefulness of ML potentials in modelling real GST-based devices.

Chapter 6 demonstrates the full-loop switching process of device-scale GST models, moving one step further to simulate nanosecond-scale operations under real device conditions. To this aim, a new ML potential is developed based on the previous training database of the GAP potential for GST. Notably, the hyperparameter optimisation will be shown, which enables a direct transfer from GAP models to ACE models. The further increase in computational efficiency and the quantum mechanical accuracy of ACE models will then be demonstrated. Lastly, I will show how the ACE potential for GST can be used in real device scale MLMD simulations to model the realistic RESET and SET operations in both confined cross-bar arrays and mushroom-type devices.

Chapter 7 summarises the conclusions of this thesis and provides an outlook of ML-based atomistic modelling for real device simulations of PCMs in the future.

Chapter 3

Methods

This chapter provides necessary background and technical details used in this thesis. In particular, this chapter focuses on the introduction of methods used to create machine-learned interatomic potentials (Section 3.1), including generation of training database (Section 3.2) and optimisation of hyperparameters (Section 3.3). An example that investigated the structural properties of amorphous red phosphorus is given (Section 3.4), demonstrating the usefulness of ML potentials in the study of complex solids. Lastly, the quantification of per-atom crystallinity using similarity-based measurement is discussed (Section 3.5), which is crucial to better study the crystallisation process of PCMs in the following chapters.

3.1 Machine-learned interatomic potentials

As mentioned in the previous section, atomistic simulations (in particular, molecular dynamics simulations) are of great importance to study microscopic mechanisms of PCMs. The reliability of atomistic simulations is largely dependent on the accuracy of modelling the target potential energy surface (PES). In chemistry and materials, first-principle calculations provide quantum-mechanical accuracy across the periodic table, which serve as the flagship method in describing PES of complex phases in PCMs, and become a ground-truth reference in computational chemistry and materials. The accuracy of AIMD simulations lends itself well to simulate a wide range of PCMs; however, due to the heavy costs of AIMD simulations, which scale approximately as N^3 (N is the total number of atoms) with the system sizes, the scope of these studies has been largely restricted in length and timescale (Fig. 3.1).

By contrast, empirical potentials parametrise the PES as a function of atomic degrees of freedom (e.g., atomic positions, bond lengths, or bond angles) usually based on classical equations of physics or chemistry. For example, the embedded

atom model (EAM), a many-body interatomic potential, expresses the energy as a sum of functions of the separation between an atom and its neighbours [159, 160]. The EAM contains two parts: the first part is a pair potential between two atoms representing the electrostatic core–core repulsion and the second part an embedding term representing the energy required to place an atom in the electron density. Note that empirical parameters in empirical potentials have to be carefully selected or fitted based on experimental data and/or high-level quantum mechanical calculations. Due to the absence of complex electronic property calculations in the inference process, the computational efficiency of empirical potentials is much improved as compared to AIMD (Fig. 3.1). However, very few empirical potentials were used in the theoretical study of PCMs, because they lack the accuracy and transferability (i.e., the ability to extrapolate unseen structures) to describe the chemical (e.g., multiple elements) and structural (e.g., phase transition process) complexity in PCMs.

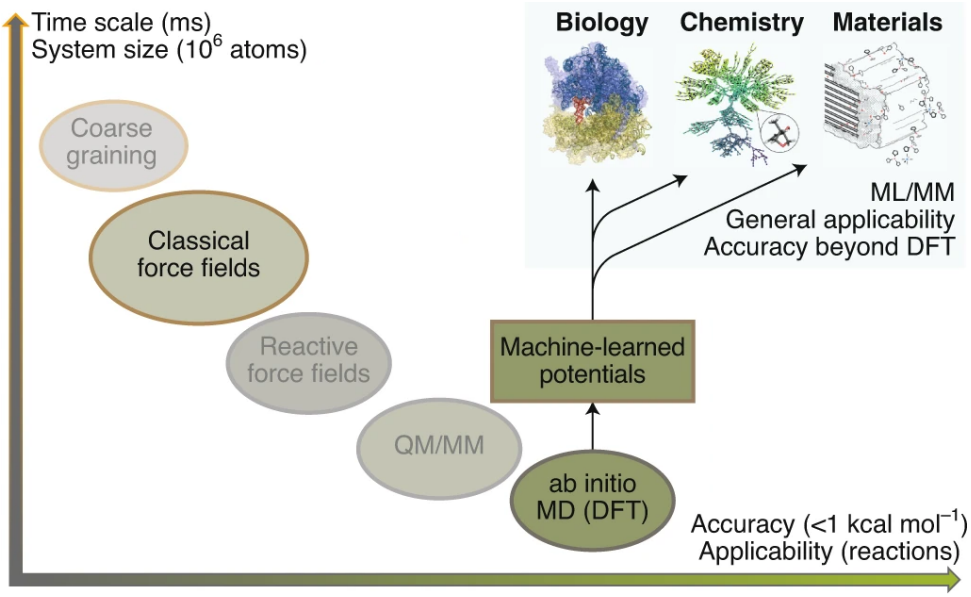


Figure 3.1: Overview of various atomistic simulations, showing a trade-off between the computational cost and the accuracy of atomistic simulations. Here, the comparison of ab initio MD, classical force fields, and ML potentials is highlighted. Potential future applications are shown in the blue box (highlighting the use of ML potentials in biology, chemistry, and materials). The "Coarse graining", "Reactive force fields", and "QM/MM" boxes are faded out as these methods are not discussed in this section. The figure is reproduced with permission from Ref. 161. Copyright 2021, Springer Nature.

"Machine-learned" potentials are fitted to a suitably chosen reference set of quantum-mechanical data, and so create a mathematical model of the PES that describes the interactions between atoms [162–166]. Once an ML potential has been developed and carefully validated, it enables simulations that are several orders of magnitude faster than those with AIMD. More importantly, ML potentials are more transferable than the empirical potentials in complex material systems, enabling accurate simulations of chemically and structurally complex compounds [118]. Hence, ML potentials are beginning to be applied to a range of challenging materials science research questions, such as the modelling of phase-change memory materials [119, 121, 137], catalysts [167], or battery materials [168].

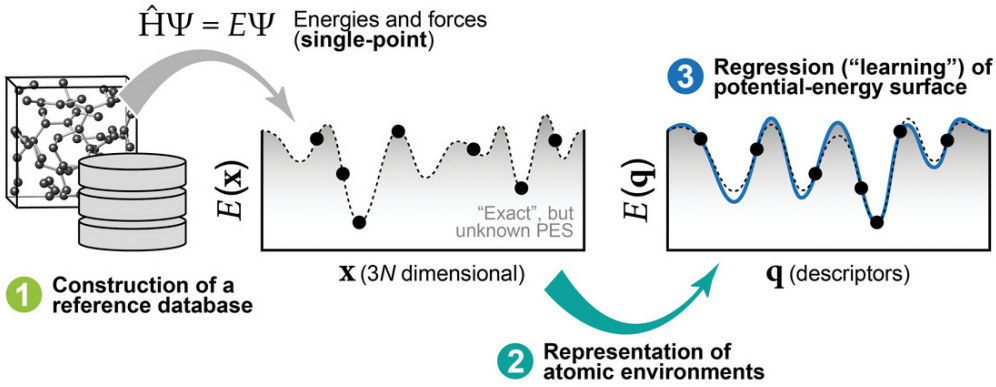


Figure 3.2: Machine-learning methodology for materials simulations. General overview of how ML-based interatomic potentials are constructed: from left to right, assembling a database of representative structural models, computing energies and forces using a reference quantum-mechanical method (typically, DFT), expressing the atomic structure in "machine-readable" form using descriptors, and finally regressing ("learning") the potential-energy surface. The figure is adapted with permission from Ref. 118. Copyright 2019 Wiley-VCH.

Developing an ML potential can be divided into three steps [118]. The first step is to generate a robust reference database. To cope with large chemical complexity in PCMs, this database has to be as complete and diverse as possible. In this step, highly converged DFT calculations are performed to obtain per-structure energies and per-atom forces for all training structures, a process also known as "database labelling". More details regarding the strategies to generate a good reference database for PCMs will be given in Section 3.2.

With a training database ready in hand, the second step is to quantitatively describe the local structural environments of each structure in the reference database using mathematical descriptors. These descriptors, which are centred on a given atom, contain the structural information of neighbouring environments and must satisfy various symmetry requirements, including permutational, translational, and rotational invariance [169]. The degree of locality of a descriptor can be controlled via adjusting a cutoff (i.e., a radius of the local environment). Despite the large flexibility to use different descriptors with different ML regression algorithms, the state-of-the-art ML potentials have preferred descriptors. For example, Gaussian approximation potential (GAP) based on the smooth overlap of atomic positions (SOAP) descriptor has been widely used in different complex material systems [164, 170, 171]. Popular neural-network-based potentials prefer the implementation of Behler's atom-centred symmetry functions (ASCFs) [120]. In the recently developed atomic cluster expansion (ACE) framework, atomic properties are constructed for each local atomic environment, leading to fast and also accurate inference [172].

All of the descriptors mentioned above depict the locality of a neighbour structural environment defined by a fixed radius cutoff (Fig. 3.3a). In addition to these locality-based structural descriptors, more recently, graph description together with the implementation of message-passing neural networks has been introduced. These graph-based descriptions have relatively smaller radius cutoffs as compared to the locality-based descriptors. Instead, the range of atomic interactions in the training structures can be extended via a number of message-passing steps (Fig. 3.3b).

The third step of potential training is to perform regression tasks so as to fit a PES for a given material system. As mentioned above, most ML regression algorithms have their preference for specific local structural descriptors or graph descriptions. In Ref. 141, ML potentials based on local structural descriptors are called local ML potentials, whereas ML potentials using graph description and message-passing neural networks are called graph ML potentials (Fig. 3.3). A key advantage of

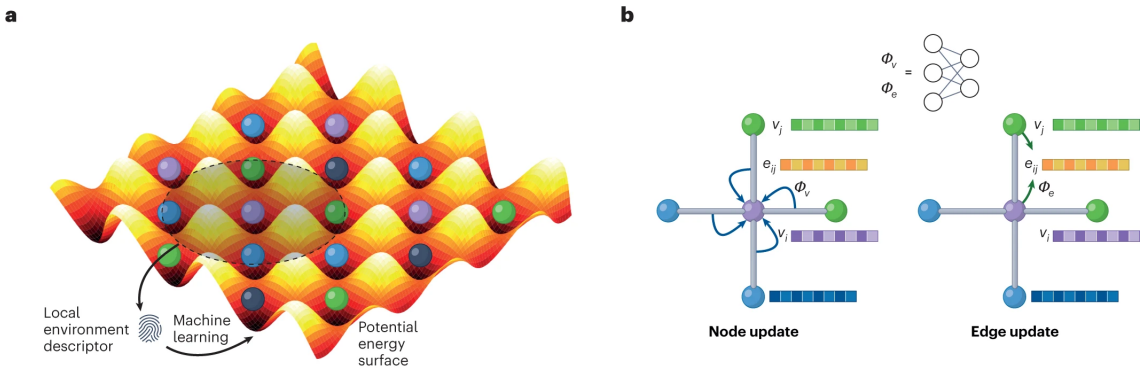


Figure 3.3: Two different ways to describe local environments in ML potentials. (a) Locality-based structural descriptors encoded the local environments of a fixed radius cutoff via atomic density, bond distances and angles, and so on. The encoded local environments are then used to fit the potential energy surface of a given system. (b) Graph-based description is combined with message-passing neural networks to describe local structural features. Atoms and interatomic bonds are regarded as nodes and edges in a graph. Structural information can be updated and flow through the graph via message passing (arrows). The figure is taken with permission from Ref. 141. Copyright 2023, Springer Nature America, Inc.

local ML potentials is their computational efficiency: for example, MD simulations using ACE potentials only take one day to run millions of MD steps for a structural model of millions of atoms over several nanoseconds [173, 174]. However, local ML potentials suffer from a combinatorial explosion in the number of features when the number of elements in the system largely increases (e.g., in multi-component systems). More training data have to be fed to the potentials to enable an accurate description of the chemically complex system. Hence, local ML potentials are mostly used in chemically simple material systems.

By contrast, graph ML potentials do not suffer from the combinatorial explosion in the number of features with the number of distinct elements in multi-component systems, due to the use of message-passing neural networks to dynamically learn the feature representations of atoms. Instead of defining static features for every possible interaction between elements (in the conventional local descriptor-based ML potentials), graph-based methods dynamically learn universal interaction patterns using adaptive features from the atomic structure and their neighbours through the graph structure (in which atoms are defined as nodes and bonds are defined

as edges). This bypasses the need for manually enumerating all possible interactions, and drastically reduces the feature space as the number of elements increases, thus making graph-based ML potentials more data-efficient, comparing to local ML potentials. However, the message-passing networks are computationally heavier as compared to the inference process in local ML potentials, because the former needs to extensively propagate the dynamically updated features across the graph (i.e., the message-passing process). Hence, graph ML potentials are more computationally expensive to run [175].

Given the difference in computational efficiency between the local and graph ML potentials, the former becomes a better choice for large-scale atomistic simulations in PCMs. Hence, this thesis focuses on local ML potentials, and all ML potentials discussed in the following are referred to as local ML potentials. The details of two local ML potentials used in this thesis, GAP and ACE, are given respectively in the sections below.

Gaussian approximation potential

Gaussian approximation potential framework has been used in the study of many material systems [176–180]. In fact, the name of "GAP" partly comes from the implementation of the Gaussian process (GP) regression, a kernel-based regression method, to construct the relationship between the atomic structural information and potential energy surface. Detailed introductions regarding GAP and GP are given in a review by Deringer et al. [171].

A relatively simple way to understand GP from a weight-space view is to define a linear combination of M basis functions:

$$\tilde{y}(\mathbf{x}) = \sum_{m=1}^{M} c_m k(\mathbf{x}, \mathbf{x}_m) \tag{3.1}$$

where the estimated value $\tilde{y}(\mathbf{x})$ is used to approximate the true value $y(x)$. The basis functions, k, are located in the input space, $\mathbf{x}_m$, comprising the representative set,

$\{\mathbf{x}_m\}_{m=1}^{M}$. The pre-factors c_m are coefficients. Note that $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ represents the similarity between the functions at two arbitrary locations, $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$. The function is symmetric to swapping its arguments, viz. $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}', \mathbf{x})$, and it is positive semidefinite.

The fitting of the GP regression model to the known data points (i.e., the input training data points) is completed by finding the coefficients $\mathbf{c} = (c_1, ..., c_M)$ that minimise the loss function,

$$\mathcal{L} = \sum_{n=1}^{N} \frac{[y_n - \tilde{y}(\mathbf{x}_n)]^2}{\sigma_n^2} + R \tag{3.2}$$

.

Two terms are shown in the loss function. The first term measures the difference between the true value and the estimated value. However, minimising the first term alone will cause overfitting of the GP regression model due to the largely flexible functional form. The second term R is a regularisation term, and the Tikhonov regularisation is used, which is defined as

$$R = \sum_{m,m'}^{M} c_m k(\mathbf{x}_m, \mathbf{x}_{m'}) c_{m'} \tag{3.3}$$

.

The parameters σ_n control the relative importance of individual data points, which adjust the balance between the accurate description of the fitting data and the overall smoothness of the model.

If using a uniform $\sigma_n = \sigma$ parameter for all fitting data points, σ_n then represents a regularisation "strength". Hence, the loss function can be written in matrix form:

$$\mathcal{L} = (\mathbf{y} - \mathbf{K}_{NM}\mathbf{c})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{K}_{NM}\mathbf{c}) + \mathbf{c}^{\mathrm{T}} \mathbf{K}_{MM} \mathbf{c} \tag{3.4}$$

.

The matrix elements are defined as:

$$[\mathbf{K}_{NM}]_{nm} = k(\mathbf{x}_n, \mathbf{x}_m) \tag{3.5}$$

and $\mathbf{y} = (y_1, ..., y_N)$ is the true target values. N is the total number of fitting data and M is the number for the representative data points. Because the kernel function is symmetric, then $\mathbf{K}_{NM}^{\mathrm{T}} \equiv \mathbf{K}_{MN}$. $\boldsymbol{\Sigma}$ is a diagonal matrix of size N, comprising all σ_n values, with $\boldsymbol{\Sigma}_{nn} = \sigma_n^2$. To minimise the loss function, solutions are searched that satisfy

$$\nabla_{\mathfrak{c}^{\mathrm{T}}} \mathcal{L} = \mathbf{0} \tag{3.6}$$

and thus

$$-\mathbf{K}_{MN}\boldsymbol{\Sigma}^{-1}\mathbf{y} + \mathbf{K}_{MN}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{NM}\mathbf{c} + \mathbf{K}_{MM}\mathbf{c} = \mathbf{0} \tag{3.7}$$

.

Hence, the coefficients $\boldsymbol{c}$ have an analytical expression:

$$\mathfrak{c} = (\mathbf{K}_{MM} + \mathbf{K}_{MN}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{NM})^{-1}\mathbf{K}_{MN}\boldsymbol{\Sigma}^{-1}\mathbf{y} \tag{3.8}$$

and once all fitting coefficients are determined, the GP regression model can then be used to predict based on a new location $\mathbf{x}$ using Eq. 3.1.

Note that if the total number of input data points is the same with the number of representative data points, i.e., $N = M$, then a "full GP regression" model is obtained. Otherwise it is a "sparse GP regression" (i.e., $N > M$).

When it comes to GAP, the total energy of a given structure is decomposed into per-atom energy, ε_i:

$$E_{\mathrm{total}} = \sum_{i}^{\mathrm{atoms}} \varepsilon_i = \sum_{i}^{\mathrm{atoms}} \varepsilon(\{\mathbf{r}_{ij}\}), \tag{3.9}$$

where $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ indicates the relative position between the centred atom i and its neighbour atom j in a local environment defined by a fixed radius cutoff, r_c. To use GP formalism to fit a GAP potential, however, the estimated values, i.e.,

atomic energies, cannot be directly obtained from DFT calculations that only yield the total energy of a given structure. Atomic forces and stresses are also inputted in the fitting of GAP, which are derivatives of the total energy with respect to the atomic positions and the lattice deformation, respectively. Hence, it is important to consider a GP regression task that estimates a function when it is not possible to directly have access to the values of the function but only to its derived properties. In Refs. [164, 181], a formalism has been introduced to solve this challenge using GP regression. To focus more on the GAP itself, this part of discussion is skipped.

To encode the local structural environments, the smooth overlap of atomic positions (SOAP) descriptor is used representing a many-body (i.e., more than three-body) interaction (Fig. 3.2). To obtain the SOAP descriptor of the neighbourhood of a given atom i, a set of atomic densities are built:

$$\rho^{i,a}(\mathbf{r}) = \sum_j \delta_{aa_j} \exp\left[\frac{-|\mathbf{r}-\mathbf{r}_{ij}|^2}{2\sigma_a^2}\right] f_{\text{cut}}(r_{ij}) \tag{3.10}$$

where the sum runs over all neighbour atoms j of element a and the neighbourhood is defined as a sphere of a cutoff r_c. $f_{\text{cut}}(r_{ij})$ is a cutoff function that smoothly goes to 0 at r_c. The smoothness of the kernel is controlled by σ_a. To ensure rotational invariance required in the descriptor for any ML potential, the atomic density is expanded in a basis of orthogonal radial functions, $R_n(\bar{r})$, and spherical harmonics, $Y_l^m(\hat{\mathbf{r}})$:

$$\rho^{i,a}(\mathbf{r}) = \sum_{nlm} c_{nlm}^{i,a} R_n(r) Y_l^m(\hat{\mathbf{r}}) \tag{3.11}$$

where the combination coefficients $c_{nlm}^{i,a}$ are the expansion coefficients and can be expressed as:

$$c_{nlm}^{i,a} = \int \mathrm{d}\mathbf{r} R_n(r)^* Y_l^m(\hat{\mathbf{r}})^* \rho^{i,a}(\mathbf{r}) \tag{3.12}$$

with n and l as convergence parameters running up to a suitable maximum.

By summing the atomic densities over all atoms in the neighbour environment (namely, via a symmetrised combination of the expansion coefficients), an analytical

form of the SOAP kernel (i.e., the rotationally invariant power spectrum) is given:

$$p_{nn'l}^{i,aa'} = \frac{1}{\sqrt{2l+1}} \sum_{m} (c_{nlm}^{i,a})^* c_{n'lm}^{i,a'} \tag{3.13}$$

where the SOAP power spectrum is normalised to yield a unit length vector. Note that the SOAP vector is smooth and continuous with respect to atomic displacements and invariant with respect to physical symmetries. It only contains two free parameters, a cutoff r_c that defines the local environment and the smooth value σ_a. To implement the SOAP descriptor into the GP regression model, a kernel function should be defined which allows to compare the similarity between two atomic environments, denoted A and A'. Instead of using a standard Gaussian kernel, most up-to-date applications use low-order polynomial kernels:

$$k(A, A') = (\xi \cdot \xi')^{\zeta} \tag{3.14}$$

where ζ determines the order of the polynomial kernels. Increasing the ζ value will not increase the computational cost. ξ and ξ' are two feature vectors corresponding to the normalised SOAP power spectrum vectors, viz. $\xi = \mathbf{p}/|\mathbf{p}|$, and the power spectrum $\mathbf{p}$ of atom i contains the components defined in Eq. 3.13, viz.

$$\mathbf{p}_i = \{p_{nn'l}^{i,aa'}\} \tag{3.15}$$

However, fitting a GAP potential to only the many-body SOAP descriptor might not always yield optimal results. The resultant GAP model fails to learn the basic two-body interaction term: the repulsive interatomic interaction at a very small atomic distance (e.g., less than 1 Å for carbon dimers [176]) and the attractive regime of interatomic bond formation on the length scale of several Ångströms. In most cases in bulk solid, very few dimers (i.e., two atoms in a large vacant area) can be found, and hence the repulsive force (energy scale of tens of electronvolts and

higher) between any two atoms is of more importance to be considered. The lack of this repulsive term will result in a severe atomic clustering in the practical MD simulations. In addition, sometimes, the three-body term, i.e., interactions among any given three atoms, is also important, e.g., in carbon [176], and thus is considered as part of the contribution to the energy.

To consider the contributions from the many-body and low-body-order (2- and 3-body) terms in the total energy of a given atom, suitable descriptors and kernels should be defined for each term and used in the the "linear functional observations" framework. For the pair interaction, the distance between two atoms is the widely-used choice. For the three-body term, either two distances and an angle or three distances are suitable. Note that sometimes the three-body term is approximated to depend only on the angle between three atoms, to reduce the number of parameters. For the many-body term, the SOAP descriptor is constructed. In practice, the combination of two-body and many-body contributions can already well describe the system: the former takes care of the basic repulsive interparticle forces and the latter focuses on complex many-body interactions including bonds between dimers and angles between triplets, which serves as a good balance between accuracy and efficiency. In fact, the many-body SOAP descriptor has already included the description of two- and three-body interactions, but is not that efficient due to its intrinsically high-dimensional nature. Hence, the total energy expression of a combination of two- and many-body SOAP can be written as:

$$\begin{aligned} \mathrm{E} = &\sum_{ij} (\delta_2)^2 \sum_{m=1}^{M_2} c_{2,m} \exp\left[-\frac{|r_{ij} - r_m|^2}{2\theta^2}\right] \\ &+ \sum_{i} (\delta_{\mathrm{MB}})^2 \sum_{m=1}^{\mathrm{M_{MB}}} c_{\mathrm{MB},m} k(\xi_i, \xi_m) \end{aligned} \tag{3.16}$$

where the weights δ_2 and δ_{MB} control the relative contributions of the different terms and have units of energy. In particular, the two-body term uses a covariance function of squared exponential kernels (Gaussian kernels), representing a one-dimensional

sparse GP using M_2 representative points located at a regular grip of interparticle distances r_m. In contrast, the many-body SOAP term uses a covariance function of dot product, indicating a high-dimensional sparse GP using M_{MB} representative atomic environments from the reference database. The pre-factors of each term c are the fitting coefficients.

The fitting process of GAP potentials is implemented in the QUIP package, which can be openly available for non-commercial use at https://github.com/libAtoms/QUIP.

Atomic cluster expansion

In recent years, the atomic cluster expansion framework has shifted the Pareto front for cost and accuracy of data-driven ML interatomic potentials [182]. As compared to other mainstream ML potentials, e.g., GAP, the near-DFT accuracy and largely improved computational efficiency of ACE potentials have underpinned the development of a wide range of state-of-the-art interatomic potential models [173, 183–185]. For example, a general-purpose ACE potential for carbon [183] was developed based on a large training database outperforms previously reported models [176] in terms of both accuracy and speed. Moreover, an ACE model of Si–O binary system, very recently developed by Erhard et al., was shown to accurately describe complex nanoscale structural features in this key binary system [173].

Unlike most ML potentials, the ACE framework produces a complete set of basis functions which describe local atomic environments of training structures. ACE formulations can be defined in either a linear or a mildly non-linear way, in large contrast to other highly non-linear ML potentials, e.g., neural network potentials, which lack a clearly defined physical meaning. Detailed introductions of the ACE framework are given in Refs. [172, 174, 182], and a brief introduction is provided below.

ACE potentials are built from general atomic "properties", φ_i, which is the function of the local atomic environment of atom i. The atomic property is expanded

across body-ordered functions from the set of neighbours of each centred atom. Any given atomic property of atom i can therefore be expanded as a function of the local environment of this atom:

$$\varphi_i = \sum_{\mathbf{v}} c_{\mathbf{v}} \mathbf{B}_{i\mathbf{v}} \tag{3.17}$$

where the expansion coefficients $c_{\mathbf{v}}$ and basis functions ($\mathbf{B}_{i\mathbf{v}}$) share multi-indices $\mathbf{v}$. Note that the basis functions fulfill translational, rotational, inversional, and permutational (TRIP) invariances for the representation of scalar variables of absolute atomic positions, as well as equivariances for the expansion of vectorial or tensorial quantities that are exploited in a graph neural network-based transferable force field [143, 144]. To achieve the TRIP invariances in basis functions $\mathbf{B}_{i\mathbf{v}}$, linear combinations of permutation-invariant (but not rotational- and inversional-invariant) many-body basis functions $\mathbf{A}_{i\mathbf{v}}$ are considered:

$$\mathbf{B}_{i\mathbf{v}} = \sum_{\mathbf{v}'} \mathscr{C}_{\mathbf{v}\mathbf{v}'} \mathbf{A}_{i\mathbf{v}'} \tag{3.18}$$

where the generalized Glebsch-Gordan coefficients $\mathscr{C}_{\mathbf{v}\mathbf{v}'}$ act as a filter to remove basis functions that are not rotational- and inversional-invariant. To achieve high computational efficiency, the basis functions $\mathbf{A}_{i\mathbf{v}}$ can be constructed recursively via a product of atomic bases $A_{i\mathbf{v}'_t}$:

$$\mathbf{A}_{i\mathbf{v}} = \prod_{t=1}^{\nu} A_{i\mathbf{v}_t} \tag{3.19}$$

where the order of the product ν determines the body order of basis function $\mathbf{A}_{i\mathbf{v}}$. The atomic base $A_{i\mathbf{v}}$ is obtained via a projection of local basis functions on the atomic density:

$$A_{iv} = \langle \phi_{\mathbf{v}} | \rho_i \rangle \tag{3.20}$$

where $\phi_{\mathbf{v}}$ is the one-particle basis function and ρ_i is the atomic density of the centred atom i. They are given by

$$\rho_i = \sum_{j}^{j \neq i} \delta(r - r_{ji}) \tag{3.21}$$

$$\phi_{\mathbf{v}} = R_l^n(r_{ji}) Y_l^m(\hat{\mathbf{r}}_{ji}) \tag{3.22}$$

where radial function $R_l^n(r_{ji})$ and spherical harmonics $Y_l^m(\hat{\mathbf{r}}_{ji})$ are used, similar with that defined in SOAP (Eq. 3.11). Hence, the atomic base $A_{i\mathbf{v}}$ can be re-written as

$$A_{iv} = \sum_{j}^{j \neq i} \delta(r - r_{ji}) R_l^n(r_{ji}) Y_l^m(\hat{\mathbf{r}}_{ji}) \tag{3.23}$$

Note that r_{ji} is the distance between the neighbour atom j and the centred atom i ($j \neq i$), and $\hat{\mathbf{r}}_{ji}$ represents the direction. The index $\mathbf{v} = (nlm)$ is cumulative, and n differentiates between orbitals with the same angular quantum numbers l and m.

The potential energy of the i-th atom is described as a combination of several atomic properties defined in Eq. 3.17. In the simplest case, the atomic energy, ε_i, is evaluated in a linear way:

$$\varepsilon_i = \varphi_i^{(1)}, \tag{3.24}$$

However, ACE potentials are not limited to using a single atomic property: introducing non-linear behaviour through a Finnis–Sinclair-like embedding, for example, results in

$$\varepsilon_i = \varphi_i^{(1)} + \sqrt{\varphi_i^{(2)}} \tag{3.25}$$

where the two properties are generalizations of the pairwise repulsion and density of the widely used Finnis–Sinclair potential [186]. The per-atom energy can be further extended to an arbitrary number of P atomic properties ($p = 1, \ldots, P$) which all enter as arguments into a non-linear function, denoted as F:

$$\varepsilon_i = F(\varphi_i^{(1)}, \ldots, \varphi_i^{(P)}) \tag{3.26}$$

where i is the atomic index and . In Ref. 173, Erhard et al. reported that a complex functional form ($P = 8$) outperformed linear and Finnis–Sinclair type embedding of ACE models for the Si–O system while only leading to a $\approx$ 15% increase in computational requirements. A full systematic study of the number, exponents, etc. of the terms that enter F is yet to be reported.

The fitting of ACE potential is carried out using the software package pacemaker [174], which can be openly available for non-commercial use at https://github.com/ICAMS/python-ace. A detailed discussion regarding the hyperparameter optimisation of ACE models is given in Section 3.3.

3.2 Generation of training database

As mentioned in the previous section, three important steps are required to fit ML potentials: generation of training data, implementation of structural descriptors, and regression using ML algorithms [118]. In fact, some structural descriptors are sometimes coupled with specific ML potential frameworks, e.g., the smooth overlap atomic position (SOAP) is mostly used together with Gaussian process method. The quality of any ML potential largely relies on two factors: (1) the functional form of the ML potential models; and (2) the training data on which the model is developed [187]. The former has attracted abundant interests in the field of modern atomistic simulations, and innovation for architectures of ML potentials has been increasingly accelerating over the past two decades, giving rise to a wide range of choices (as discussed in Section 3.1) [141, 171, 175, 188, 189]. However, the training data also play a non-negligible role, where it determines the best accuracy that an ML model can approach, i.e., the upper bound performance of ML potentials [187].

A consensus in ML potential training is "garbage in, garbage out" – a badly constructed reference database will lead to failure in accurately describing the PES of desired material systems. From a general perspective, a supervised atomistic ML model establishes the mapping between a set of input features x and the desired properties $\hat{y}$ [187]. A training database contains a variety of structural configura-

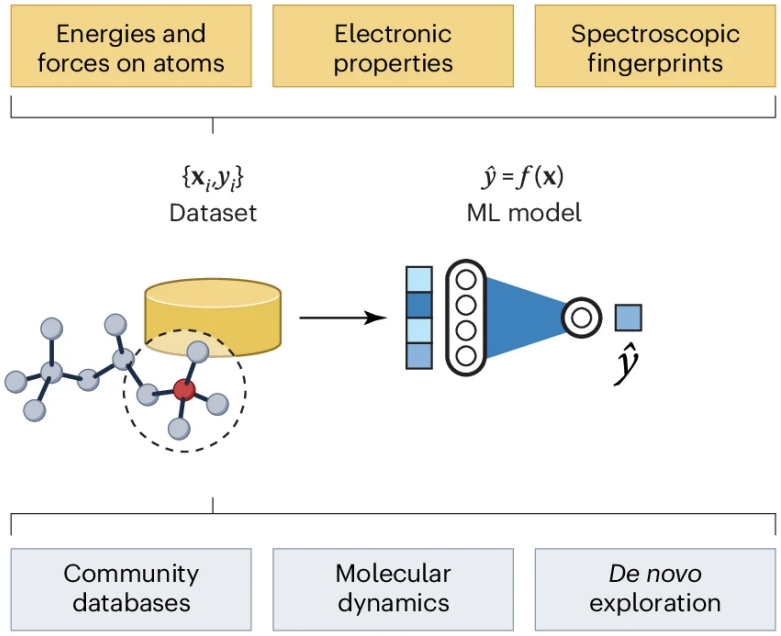


Figure 3.4: A typical workflow for an atomistic ML model that establishes the relationship between a set of features, x, and the corresponding model predictions, $\hat{y}$. The dataset contains data locations, and the target (desired) properties. Training structures can be collected via different approaches, including publicly available databases, molecular dynamics sampling, or de novo exploration. Labels are typically obtained from quantum-mechanical computations. The dataset is indicated in yellow, and the featurisation and regression tasks are shown in blue. The figure was taken with permission from Ref. 187. Copyright 2024, Springer Nature America, Inc.

tions, including the locations of atoms (x_i), per-structure energies (E_i), and per-atom forces ($\vec{F}_i$). Energies and forces are computed using DFT as separate labels, serving as target functions for ML models to learn. Such training database is categorised as a "single-purpose" database, which mainly aims at one specific fitting task, such as the reconstruction of PES of a given material system, a prediction of the density of states of complex solids, or atomic charges and dipoles required for spectroscopy prediction (Fig. 3.4).

Figure 3.4 also shows typical approaches to construct a training database. A good starting point is to collect training data from existing community databases, which contain computational and experimental domain knowledge, including openly available high-throughput databases (e.g., QM7 database for organic molecules [190, 191]) and popular crystal structure databases (e.g., the Inorganic Crystal Structure

Database and the Materials Project). Sufficient labelled data can be taken directly to set up an initial training database for further iterations. Based on the off-the-shelf data, multiple rounds of iterative training can be carried out, for example, via molecular dynamics simulations or Boltzmann generators [192]. Additional data can be gradually added back to the training database based on such physically guided sampling to extend the exploration of PES. Unlike the requirement of deep domain knowledge of physically-related sampling, the random structure search protocol tries to sample the PES in a de novo ("from scratch") fashion. Initially rooted in crystal structure prediction, RSS protocol has been proved to enable a wide range of sampling in complex materials systems, e.g., the high-pressure phases of matter [193], and largely increase the robustness and extrapolation ability of ML potentials in dealing with unknown phases.

Despite that collecting data for single-purpose database is becoming a routine task for many materials, it is still challenging to efficiently search for a complete set of training database for complicated systems, e.g., phase transition processes of inorganic solids or reaction paths of organic molecules. In particular, learning the phase-change process requires comprehensive descriptions of crystalline and amorphous structures of PCM, as well as the crystalline-amorphous interfaces in PCM solids. Given the large structural and chemical complexity in PCM, collecting a diverse training database takes a lot of efforts and real-world time. To address this challenge, two useful protocols will be discussed in the following, including domain-specific iterative MD sampling and de novo random structure search.

3.2.1 Iterative MD sampling based on domain knowledge

Iterative training allows for efficient reference data generation, and has become one of the necessary components in most ML potential training nowadays. It can sometimes behave like a self-correction process, especially when certain unphysical configurations, e.g., atomic clustering, are fed back to the ML model, correcting the previously incorrect behaviours of the model. In fact, iterative sampling could be

used together with many other protocols, including the RSS process which will be discussed in Section 3.2.2. Such iterative generation of training database is referred to as "standard iteration" as it emphasises the iterative process itself instead of the specific structures to sample - that is, a constant MD protocol is used throughout the iterative process. Running conventional MD simulations (e.g., a melt-quench MD simulation to create amorphous solid) using an existing ML potential and then collecting new structures for a new version of ML potential thus became the golden standard in fitting ML potentials.

The first ML potential for GeTe, which was fitted using a neural network framework, followed standard iterative MD processes to collect training data [119]. Using a early version of NN potentials, melt-quench simulations at different pressures were carried out to generate new crystalline, liquid, and amorphous structures of GeTe. Together with the snapshots taken from crystallisation simulations using AIMD, a comprehensive training database was generated, leading to an accurate NN potential for GeTe. Such ML-driven melt-quench MD simulations have also been used in collecting training databases for other material systems, including tungsten [194], carbon [176], silicon [177], phosphorus [178], and so on.

By contrast, some configurations can barely be sampled within a standard melt-quench iterative MD simulations. An example shown by Deringer et al. highlights the domain-specific iterative MD sampling for a GAP potential of carbon [176]. Standard iterative MD sampling was first performed to collect different phases of carbon. To add amorphous surfaces to the training set, melt-quenched snapshots were obtained from GAP-MD simulations. A vacuum slab was then added on the top of the melt-quenched bulk structure, which created an amorphous slab and eliminated the periodic condition along the vertical direction (Fig. 3.5a). The amorphous slab was further heated up to elevated temperatures (e.g., several thousand K) for further equilibration. Snapshots were then taken and labelled as new training data for a new version of GAP potential. It has been shown in Ref. 195 and 196 that

this C-GAP potential, developed via a domain-specific iterative MD sampling, can be used to study the the microscopic mechanism of as-deposit growth process in carbon and the formation of different types of local environments (e.g., sp^2 and sp^3) in amorphous carbon.

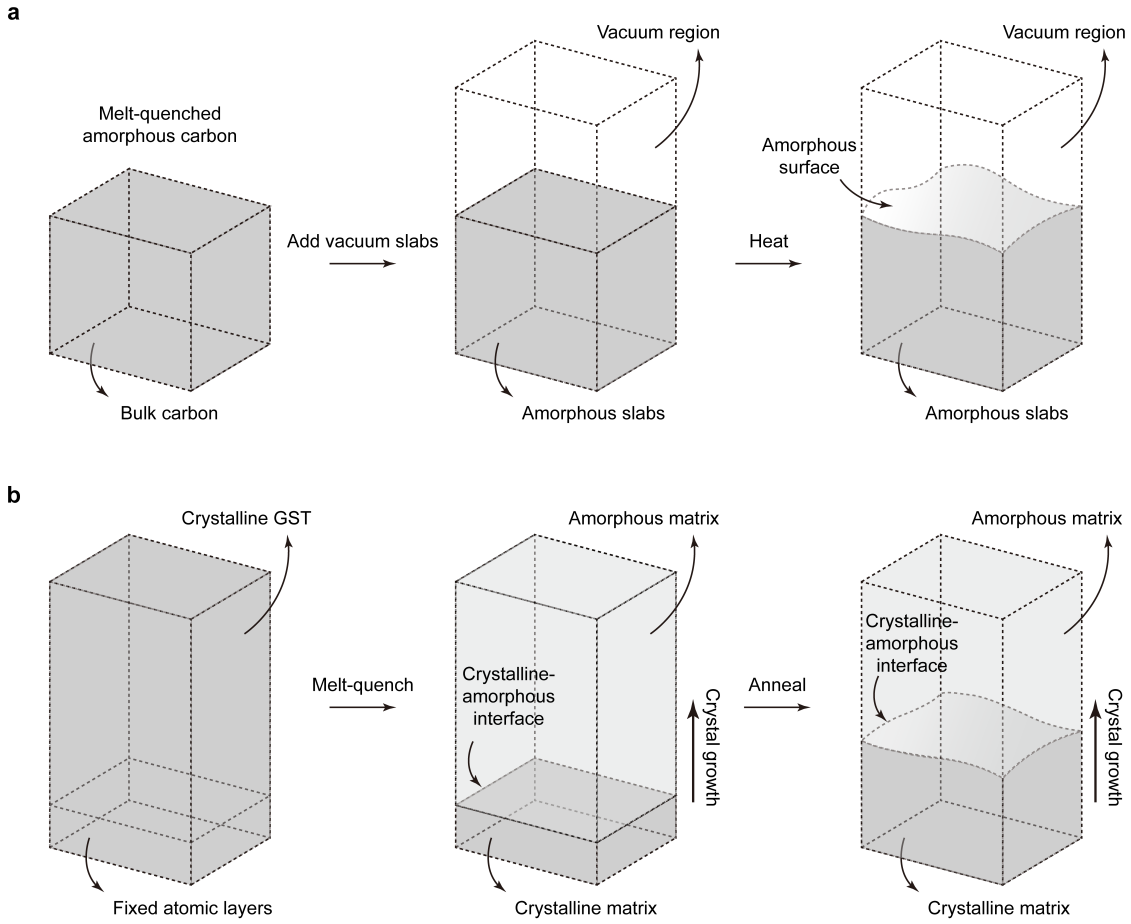


Figure 3.5: Sketches showing domain-specific iterative MD sampling protocols. (a) Domain-specific protocol to create representative amorphous surfaces of carbon [176]. A melt-quenched simulation was first performed to create a bulk amorphous structure. Then, a vacuum slab was added on the top the bulk amorphous to create an amorphous surface. The structural model was then heated up to high temperatures, and snapshots were taken to be added back to the training database. (b) A melt-quench protocol which generates intermediate configurations during crystallisation, in which a crystalline-amorphous interface was obtained. Starting from a crystalline structure, several atomic layers are fixed. Melt-quench simulations are then performed to create amorphous structures with crystalline matrix. The templated structural models are further annealed at crystallisation temperatures where crystal growth is observed. Representative snapshots of crystallisation are then taken from this process.

Domain-specific iterations are also of great importance when developing ML potentials for PCMs. For example, the crystallisation of PCMs usually takes several

to hundreds of nanoseconds to finish. In particular, nucleation has a stochastic nature causing a long incubation time [108, 110]. It thus becomes infeasible to run a MD simulation to sample the whole process of crystallisation within an acceptable real-time period. In this case, properly designed MD simulations based on deep understanding of domains enable specific learning tasks towards desired properties. In a previous study, a melt-quench protocol was carried out on a crystal structure with two fixed atomic layers, which served as a crystallisation template bypassing the stochastic incubation process (Fig. 3.5b) [109]. Following this idea, melt-quench simulations with some fixed layers can be performed to create the structural model with a crystalline-amorphous interface in a relatively small-scale cell (containing less than 200 atoms), as shown in Fig. 3.5b. Such protocol can be achieved via a proper design of MD simulations, requiring deep domain knowledge. More details regarding this melt-quench protocol to sample the intermediate configurations in crystallisation will be discussed in Chapter 4–6.

Note that iterative sampling is a general process and can be combined with other strategies, such as active learning scheme based on uncertainty indicators or query-by-committee methods, to sample the unknown structural space [173].

3.2.2 Random structure search

Unlike the domain-specific iterative MD sampling that requires laborious efforts to collect training data to tackle specific research problems, random structure search (RSS) serves as a much easier approach to explore a wide structural space during ML potential fitting. It is important that the database must contain higher energy regions of the PES, not just the local minima (e.g., well-known crystalline structures). RSS can well sample these higher-up regions in PES [197]. In fact, RSS was previously used to search experimentally unknown structures that do not exist under ambient conditions. Pickard et al. have used ab initio-driven random structure search (AIRSS) to explore high-pressure phases of complex condensed matters [198, 199]. Following this idea, the use of ML potentials has enabled a more efficient

structural search in a wide range of inorganic materials [200, 201]. More recently, RSS has been taken as the first step to include diverse structures in the training database [201–204].

The key idea of RSS is to explore configurational space through generation of random structures and subsequent structural optimisation, thus generating an accurate estimation of PES. The whole process of ML-driven RSS is carried out in an iterative way: a large number of random structures are first generated using AIRSS algorithm [198, 199]. Based on the leverage-score CUR algorithm [205], N most diverse structural models are taken as the initial training database (Fig. 3.6a). An ML potential (e.g., here a GAP potential) is then fitted to this initial database. Note that the principal component to enable a wide range of sampling of the structural space is to use a configuration-averaged kernel metric, which allows the selection of the most relevant and diverse structures at each iteration. In this case, the dissimilarity distances between candidate structures are calculated based on the smooth overlap atomic position (SOAP) method [170]. To measure the structural dissimilarity between two given structures, a configuration-averaged SOAP was used [206].

Starting from the second round, new ensembles of random structures are generated using AIRSS. Unlike the first round of data collection, these generated random structures are optimised using the GAP potential generated from the previous round. Again, N most diverse structures are selected from all the optimisation trajectories. In this process, a repetitive selection process combining Boltzmann-probability biased flat histogram sampling and leverage-score CUR is implemented (Fig. 3.6b). The former guarantees the focus on low-energy structures whereas the latter is for the selection of the most diverse structures. The selected structures are then labelled using DFT and added back to the training database. Such iterative RSS processes are repeated for several times until the resultant ML potential meets certain requirements (e.g., the required numeric accuracy or the target functional purposes).

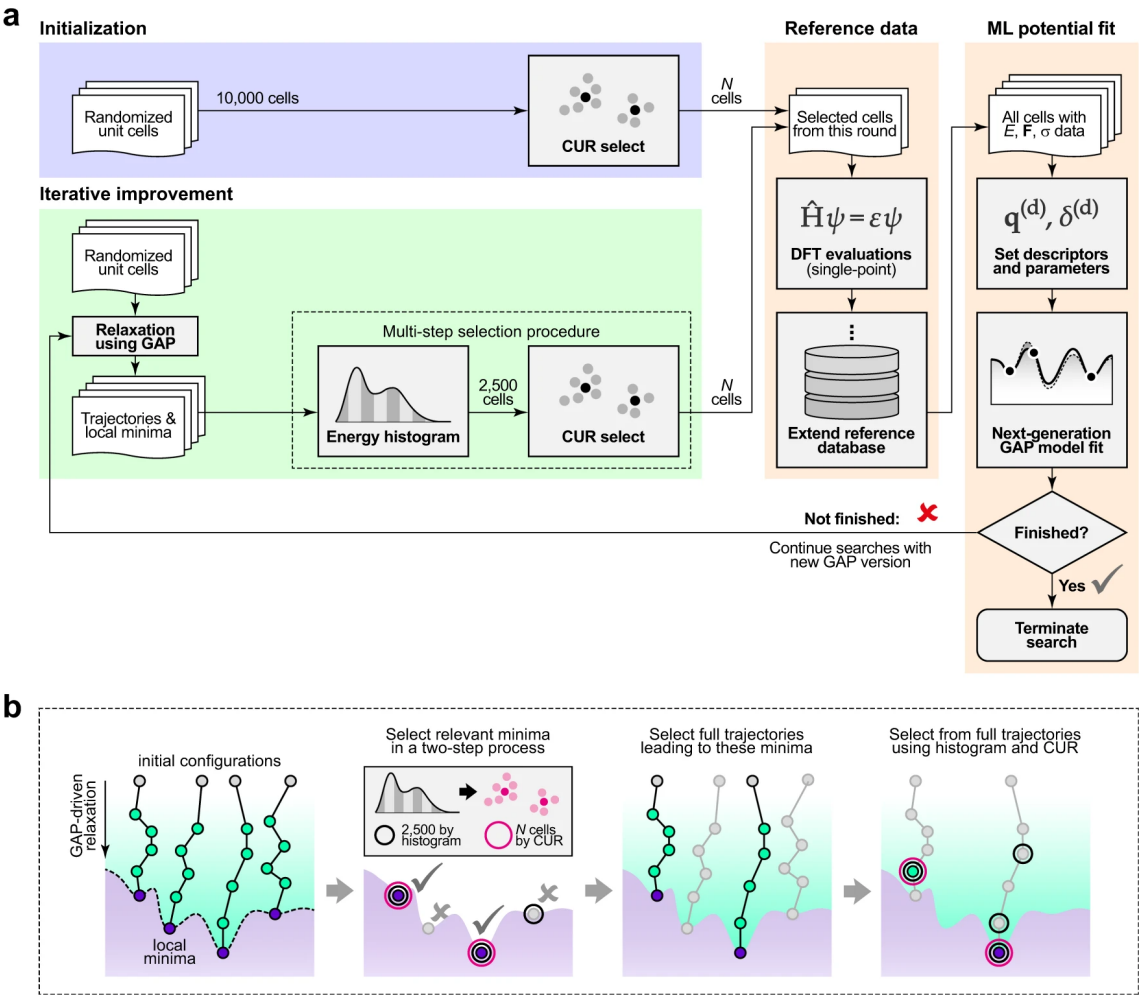


Figure 3.6: An automated protocol that iteratively explores structural space and fits machine learning (ML)-based interatomic potentials. (a) General overview of the iterative RSS sampling. From an ensemble of randomised structures (blue) generated using AIRSS, the most structurally diverse ones are selected by using the leverage-score CUR algorithm. Selected structures are labelled using single-point DFT computations and used to fit an initial GAP potential (orange). Then, new ensembles of random structures are generated. The fitted GAP potential is used to relax the new random structures (green). The most relevant snapshots are picked up. The cycle repeats until the GAP potential reaches the desired performance. (b) Illustration of the multi-step selection procedure. All optimisation trajectories are considered, sketched by connected points. A combined selection process using an energy criterion, the flattened histogram, and then a structural criterion, the CUR, is carried out. The green circles represent the intermediate points during the structural relaxation; the purple circles indicate the end points of the relaxation; and the magenta circles are the finally selected configurations that are then labelled using DFT and added back to the training database. The figure is taken with permission from Ref. 197. Original figure published under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

It has been shown that the iterative GAP-RSS process enables the collection of diverse structures and therefore produces an accurate description of PES of complex

material systems. It allows for the generation of ML potentials with minimal efforts but with efficient sampling ability. Many unknown or rarely visited structures can be searched, which goes beyond the explored configurational space of iterative MD sampling, making RSS a suitable strategy to explore the structural space in the early stages of developing ML potentials. However, note that some domain-specific configurations (e.g., the ones discussed in Fig. 3.5) can hardly be sampled by iterative RSS process. Hence, to develop a more general ML potential, which simulates largely diverse phases and complex dynamical processes (e.g., crystallisation), iterative RSS protocol can be combined with the domain-specific iterative MD sampling method. More details will be shown in Chapter 4–6 to demonstrate the use of the combination of the two sampling methods.

3.3 Hyperparameter optimisation using XPOT

Part of this section is based on a previously preprinted work on arXiv (Ref. 207), where I am one of the co-authors. Some texts and figures have been reused. I would like to acknowledge the great contributions from Daniel F. Thomas du Toit to develop the hyperparameter optimisation platform and fit various ACE potential models discussed in this thesis.

Despite the usefulness of various ML potentials in the study of chemistry and functional materials, fitting ML potentials has already posed large challenges - for example, how to efficiently collect a reference database and how to choose the hyperparameters used in the ML models? In particular, when working with emerging ML potential frameworks, a large amount of time has been used to explore feasible and reliable hyperparameters.

In fact, due to large computational costs in the process of data labelling for ML potentials (usually at the level of millions of CPU hours for a complex material system), reusing existing reference databases but with different ML fitting frameworks is becoming a wise choice for developing ML potentials and performing large-scale MD simulations with specific purposes. ML potentials lack evident

physically-motivated functional forms, serving as a "black box" in the prediction of required properties. A set of well-chosen hyperparameters, at the beginning of fitting, is of great importance to creating a robust and accurate ML potential. For mathematically complex ML models in particular, the effects of changing hyperparameters are not always immediately clear. Hence, it would be desirable to explore the space of hyperparameters automatically and to optimise them with minimal manual input.

A previous work by Thomas du Toit et al. [208] showed that hyperparameter optimisation for spectral neighbour analysis potential (SNAP; [163]) models can improve the prediction accuracy without increasing cost at runtime. By contrast, poorly chosen hyperparameters will decrease the accuracy when predicting energies and forces of complex systems [208]. The resultant hyperparameter optimisation protocol was called the cross-platform optimizer for machine-learning interatomic potentials (XPOT). In this protocol, a specific set of hyperparameters was denoted as $\mathcal{H} = \{h_1, \ldots, h_N\}$, and the space of all possible choices was taken as $\mathbb{H}$ ($\mathcal{H} \in \mathbb{H}$). Note that $\mathcal{H}$ could include discrete entries instead of continuous functions. Overall, the expectation is that hyperparameter optimization can increase the accuracy and robustness of potentials for a given reference database.

In addition to the optimisation of SNAP models, XPOT can also be used in other ML potentials with more complex mathematical forms. In a recent work [207], XPOT was introduced to fit ACE potentials based on existing reference database for Si and Sb_2Te_3 alloys.

In optimisation, users can define ranges of target hyperparameters and then XPOT automates the optimisation process by minimising a loss function iteratively. The loss function combined energy and force loss which is defined as

$$\mathcal{L} = \frac{\alpha}{[E]} \times \mathcal{L}_{\mathrm{E}} + \frac{(1-\alpha)}{[\vec{F}]} \times \mathcal{L}_{\mathrm{F}}. \tag{3.27}$$

In the above expression, square brackets represent the units, emphasizing that

the loss function, $\mathcal{L}$, is dimensionless. A single parameter α was used to control and adjust the relative weighting between the energy and force loss, so that the potential model can be optimised towards a desired characteristic.

On the one hand, the energy contribution to the loss function, $\mathcal{L}_{\mathrm{E}}$, is expressed by

$$\mathcal{L}_{\mathrm{E}} = \sqrt{\frac{1}{N_{\mathrm{cells}}} \sum_{i \in \mathbf{D}_{\mathrm{val}}} \left(\frac{\hat{E}_i - E_i}{\sqrt{n_{\mathrm{at},i}}} \right)^2} \tag{3.28}$$

where N_{cells} is the number of structures in the validation set, $\mathbf{D}_{\mathrm{val}}$, and $n_{\mathrm{at},i}$ is the number of atoms in the i-th structure from the validation set. The energy prediction of the ML model is $\hat{E}_i$, whereas the "correct" reference energy (i.e., the ground-truth) is E_i. The difference between the two is divided by $\sqrt{n_{\mathrm{at},i}}$, so as to weight the contribution of energy errors from each structure equally, regardless of the number of atoms in any given validation-set structure [209].

On the other hand, the force contribution to the loss function,

$$\mathcal{L}_{\mathrm{F}} = \sqrt{\frac{1}{N_{\mathrm{at}}} \sum_{j \in \mathbf{D}_{\mathrm{val}}} \left| \hat{\vec{F}}_j - \vec{F}_j \right|^2}, \tag{3.29}$$

is calculated by measuring the magnitude of the error between the predicted and reference force vectors. Given the forces are per-atom quantities, the overall error is simply averaged over the total number of atoms in the dataset, N_{at}.

Figure 3.7 shows the workflow of XPOT, emphasising various tasks within the optimisation loop and the relative order among them. In this loop, `PACEmaker` is one of the key components, which is used to fit ACE models. `XPOT` serves as a bridge between the potential fitting software (left-hand side in Figure 3.7) and the Bayesian optimization (BO) interface of `scikit-optimize` (right-hand side). The latter is used to optimise both discrete and continuous variables, as discussed in Ref. 208. Once users parse the manually-defined hyperparameter ranges, the programme starts to fit the potentials iteratively. Based on the results (i.e., predicted errors), new hyperparameters are chosen through BO predictions until the loss reaches the desired

value, or the maximum number of iterations is reached. A full description of the XPOT code can be found in Ref. 208.

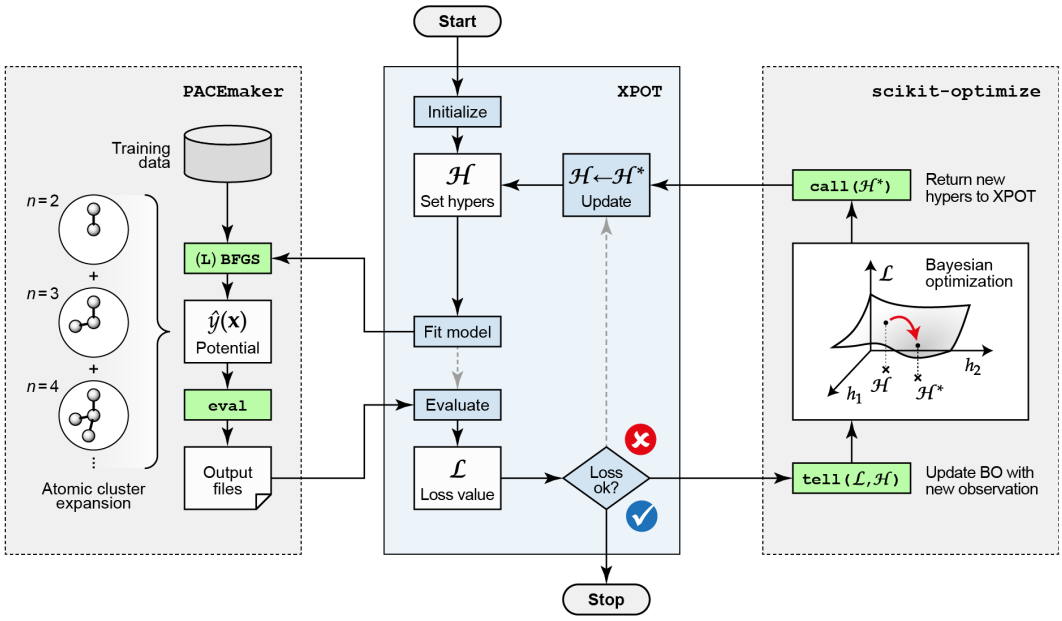


Figure 3.7: Overview of the methodology for automated optimization of ACE potentials, showing a simplified flowchart of the computational tasks involved. The core functionality of `XPOT` is highlighted in the central box in blue, and external calls to `PACEmaker` (left) and `scikit-optimize` (right) are indicated.

Note that after the optimization of XPOT, the best potential needs to be "upfitted". Upfitting continues the fitting process from an existing ACE potential, which is similar with the "fine-tuning" for neural-network-based models. As reported in Ref. 174, by tuning the relative weighting between the energy and force loss, κ, the accuracy of an ACE potential can be further improved using a two-step fitting: firstly, a large value of κ was used to converge forces (i.e., more weights on the force loss function); and secondly, a small value of κ is then used to refine the energy accuracy. For the XPOT-optimised ACE models shown in this thesis, $\kappa = 0.8$ was first used, followed by an upfitting of the same potential with $\kappa = 0.02$.

Given that the optimisation process in XPOT is agnostic to specific fitting methodologies, file parsing between XPOT and the desired fitting software can be implemented in a general fashion. For example, evaluation of the loss function in the XPOT-ACE framework is performed using the per-structure predictions pro-

vided by `PACEmaker` itself, whereas the XPOT's interface to `fitsnap` uses `LAMMPS` to drive the optimisation process. In addition, XPOT can deal with both linear and non-linear ACE models, corresponding to different P values in Equation 3.26, which provides additional flexibility to fit ACE potentials.

An example of Sb_2Te_3 is shown below to demonstrate the usefulness of XPOT when fitting an ACE potential, as a non-expert, using the previous reference data generated for GAP models [207]. The dataset used for Sb_2Te_3 is taken from the GST-GAP-22 dataset [180], which is also a largely handcrafted training database and will be discussed in greater detail in Chapter 5. As only Sb and Te were considered for the ACE model of Sb_2Te_3, all structures which include Ge were removed.

To generate the testing data for numerical validations, the original dataset was split into train and test parts in a ratio of 80:20. For each configuration type as defined in the dataset, such as "`crystalline`", "`aimd`", and "`liquid`", 80% of the configurations will be taken as the training data while the remaining 20% will be seen as the testing data. In this way, the validation data can span a wide range of structure types, and reduce the risk of overfitting the ACE model towards a single subgroup of the dataset.

An ACE potential was fitted, referred to as SbTe-XPOT-ACE, based on the aforementioned training database and via the XPOT optimisation. Again, the SbTe-XPOT-ACE showed good numerical performance as compared to the GST-GAP-22 potential [180] and a newly fitted GAP potential using the same training database of SbTe-XPOT-ACE. However, the SbTe-XPOT-ACE potential showed much increased computational efficiency (over 400× faster) in contrast to the GAP models.

A further physically-guided validation was carried out, which tested the ability to describe the complex structural features in Sb_2Te_3 using GST-GAP-22 and SbTe-XPOT-ACE potentials as compared to the AIMD reference. The structures of amorphous and liquid Sb_2Te_3 were created: a melt-quench simulation was performed on a crystalline Sb_2Te_3 model (containing 360 atoms). It was found that

the computed RDF and ADF of the liquid and amorphous Sb_2Te_3 using both SbTe-XPOT-ACE and GST-GAP-22 well matched with each other. Both of them have good agreement as compared to the AIMD reference (Fig. 3.8a–b).

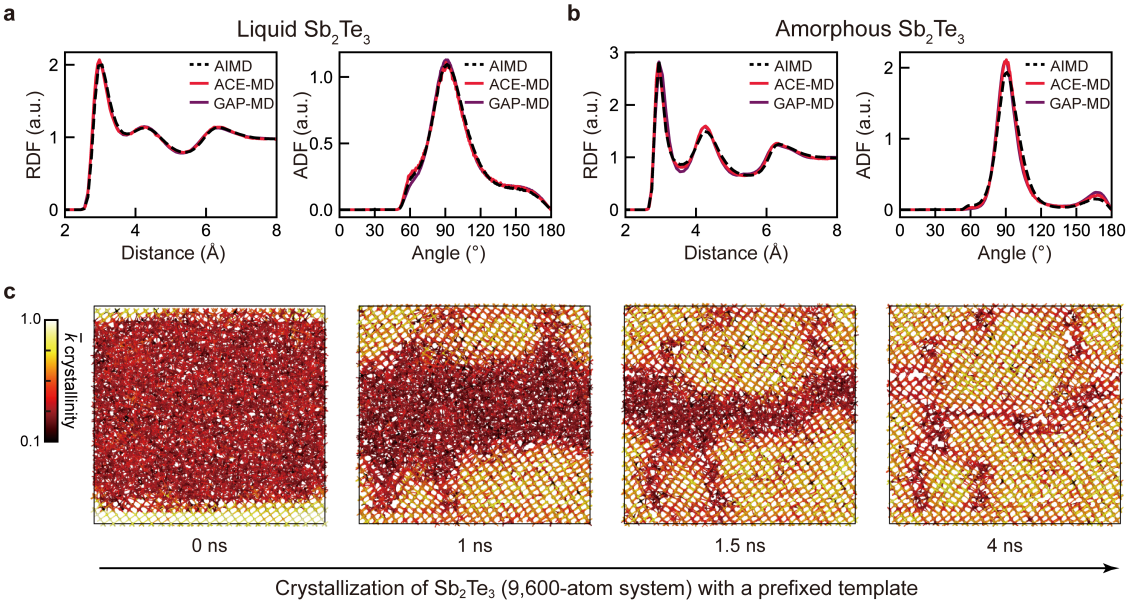


Figure 3.8: Structure and dynamics of Sb_2Te_3. (a) RDF and ADF plots for liquid and amorphous structures simulated by the SbTe-XPOT-ACE (red) and GST-GAP-22 (purple) models as well as AIMD (black). The data for the latter two are taken from Ref. 180. (b) Snapshots of an MD crystal-growth simulation. SOAP similarity is used to highlight the growth of crystalline Sb_2Te_3 (yellow) across snapshots. Several crystal grains and grain boundaries are visible.

To move beyond the basic structural validation and test SbTe-XPOT-ACE in a practical use to study the crystallisation of Sb_2Te_3, a 9,600-atom structural model of amorphous Sb_2Te_3 was created (in a box of $4.3 \times 9.0 \times 8.5$ nm^3) with a pre-fixed crystalline template to simulate the crystal-growth process. Figure 3.8c shows a crystallisation simulation for this templated structural model using the SbTe-XPOT-ACE potential. To quantify the crystallisation process, the atoms were colour coded using the SOAP-based crystallinity with respect to rocksalt-like Sb_2Te_3 to quantify the per-atom crystallinity (which is discussed in detail in Section 3.5) during the crystallisation process [114]. After the melt-quench, the templated structural model was annealed at 600 K for 4 ns, and the growth proceeded quickly at the rough crystalline–amorphous interface. Upon nanosecond crystallisation, many defects

were found (e.g., point defects and layer stacking faults) in the recrystallized model (cf. the dark red atoms in Fig. 3.8c), indicating competing growth of different crystalline regions with different crystal orientations. Notably, such local disorder is challenging to fully characterise due to the short timescales on which they happen (e.g., in the programming operations of real-world devices), and high-temperature annealing helps to eliminate the local defects, e.g., via vacancy ordering [10], resulting in an energetically more favourable crystalline phase with fewer defects [210].

To summarise briefly, hyperparameter optimization plays an important role in improving the accuracy and efficiency of desired ML potentials. Alongside dataset construction and curation [187], the choice of hyperparameters remains an important aspect of fitting robust and accurate ML models. In addition, XPOT provides an off-the-shelf tool to fit desired ML potentials based on existing reference database. In the following chapters, more details will be discussed and more examples will be shown of how ACE potentials can be developed via XPOT optimisation based on the reference databases of GAP models.

3.4 The applications of ML potentials in the study of complex solids

To showcase the capacity of ML potentials in the study of complex solids, this section shows an example of how to study the structural properties of amorphous phosphorus (a-P) using an existing general-purpose GAP potential developed for P in 2020 [178]. This GAP model was shown to be flexible enough to simulate the pressure-induced liquid–to-liquid phase transition from the molecular fluid to the network liquid, whilst also accurately describing different crystalline allotropes and the layered structure of phosphorene [178]. Now more challenging explorations of structurally complex phase, the amorphous phase, will be carried out, on which it has not yet been explicitly "trained". Part of this section is based on previously published works in Adv. Mater. 2022 (Ref. 211) and in Angew. Chem. Int. Ed.

(Ref. 212). Some texts and figures have been reused. I acknowledge the great contributions of William Kirkpatrick, Stephen R. Elliott, and Volker L. Deringer.

3.4.1 What is the structure of amorphous phosphorus?

Research interest in a-P has grown significantly because of emerging applications in batteries [213–215]. As a commercially available anode material, red phosphorus provides a large cation-storage ability with high theoretical capacities by forming binary X–P compounds (X = Li, Na, K), but it suffers from low conductivity and a large volumetric change during cycling [216]. The structure and physical properties of phosphorus are clearly important for battery applications: for example, experimental work showed that the size of a-P particles in phosphorus–carbon composite anodes has an effect on the electrochemical performance [217], and in situ transmission electron microscopy (TEM) revealed images of red phosphorus segments within a carbon nanofibre [218]. In addition, an unidentified crystalline phase of phosphorus, type-II red phosphorus, was recently discovered by various experimental tools including electron diffraction tomography, atomic-resolution scanning transmission electron microscopy (STEM), powder X-ray diffraction, and Raman spectroscopy.

Structurally ordered phases of P have been studied with DFT (see, e.g., simulations of Li adsorption on phosphorene in Ref. 219), and ab initio structure searches for Li–P and Na–P phases have been carried out [220]. However, dimensionally much more extensive structural models are likely required to correctly describe alkali-metal insertion in a-P. Realistic simulations based on ML potentials, and the resulting structural models of pristine a-P, might therefore ultimately help to understand the structural properties of a-P and control the electrochemical properties of nanostructured a-P based anodes.

What, exactly, is the structure of pristine a-P? Various models have been proposed to date. Lannin et al. suggested a quasi-two-dimensional structure with locally layer-like fragments, resembling those in black phosphorus, supported by Raman scattering and infrared absorption [221–223]. Elliott and coworkers argued

that neutron diffraction data support the presence of P8 (and possibly P9) cage-like fragments [224]; hence, the medium-range order (MRO) in a-P might be similar to that found in crystalline Hittorf's (violet) and fibrous phosphorus [225, 226]. This interpretation is consistent with early Raman spectroscopy studies, which also identified the presence of P8 and P9 cages [227, 228] and likely formation of P7 clusters [228]. By contrast, Zaug et al. carried out empirical potential structure refinement (EPSR) based on X-ray diffraction (XRD) data [229], and the resulting structural model mainly comprises P3 rings and P4 tetrahedra linked by chains of atoms. These P3 and P4 units are considerably different from the aforementioned P7, P8, and P9 ones, for which the principal structural unit is a 5-membered ring.

In addition to experimental studies, computational work provided insights into possible cluster fragments in the amorphous phase. Early AIMD simulations of P2–P8 clusters and of melt-quenched a-P led Jones and Hohl to suggest a network structure of small n-membered clusters [230, 231]. A systematic quantum-mechanical study of structural fragments was reported by Böcker and Häser [232], building on Baudler's rules for describing phosphorus structures [233]. The authors suggested that the dominant microscopic structures in a-P are tubular units and chains, consisting of clusters containing 7–10 atoms that had been discussed earlier [233, 234]. Very recently, simulated Raman spectra for crystalline allotropes and other candidate structures corroborated the absence of small chain fragments in a-P, including P4 tetrahedra [235]. Yet, all these studies are invariably limited by the system sizes that are accessible to quantum-mechanical methods.

3.4.2 Notation for cluster fragments

In the previous paragraphs, some notations (e.g., P3, P4, etc.) have already been used to represent different P clusters. Before starting discussions of amorphous phosphorus, it is thus important to define the notations of these fragments. The notations of all clusters discussed in this section are taken from Böcker and Häser's work [232], and similar notations were also used in other studies of phosphorus

clusters [230, 233]. The standard P_n notation was used for clusters consisting of n phosphorus atoms, and square brackets enclosing parts of more complex fragments (for example, "[P3]") were employed to describe the relationship between "donors and acceptors" as discussed in earlier work [232, 233]. The fragments studied here are those discussed by both theoretical and experimental work.

3.4.3 Creating a-P via a melt-quench process

Here, large-scale ML-based molecular-dynamics simulations were carried out on a timescale of nanoseconds (corresponding to millions of individual simulation steps) to study structural features and pressure-dependent changes of a-P. The melt-quench simulations were carried out to create structural models of a-P (Fig. 3.9). To obtain the LDA a-P structure, the starting configuration was taken from the high-density liquid at 1323 K at 1.01 GPa, containing 1,984 atoms in a cubic simulation cell with a density of 2.78 g cm^{-3} (green star in Fig. 3.9a). The structure was first held at 1500 K for 50 ps, so as to be strongly randomised, and then cooled to 1200 K at a rate of 10 K ps^{-1}. Note that slow cooling provides a well-relaxed (-equilibrated) structure, leading to a more ordered network containing fewer atomic defects; hence the liquid P was slowly cooled down from 1200 K to 300 K with a quench rate of 0.1 K ps^{-1}. This quench rate was also used in ML-based cooling simulations to obtain a well-relaxed structure of a-Si [236]. The obtained amorphous structure was further optimised based on the conjugate-gradient method (at fixed cell shape and volume, thereby retaining the density from the last step of the MD simulation) to obtain the final LDA phosphorus model.

The high-temperature liquid at ambient pressure consisted mainly of P4 molecules, whereas it transformed to a network liquid at higher pressure [237]. The latter was used as a starting point, decompressed, and then cooled down from that metastable liquid state: the protocol is indicated in Fig. 3.9a. A quench rate of 10^{11} K s^{-1} was used from 1200 K to 300 K – more than three orders of magnitude slower than in a previous AIMD simulation of a-P [231], and consistent with recent GAP-driven

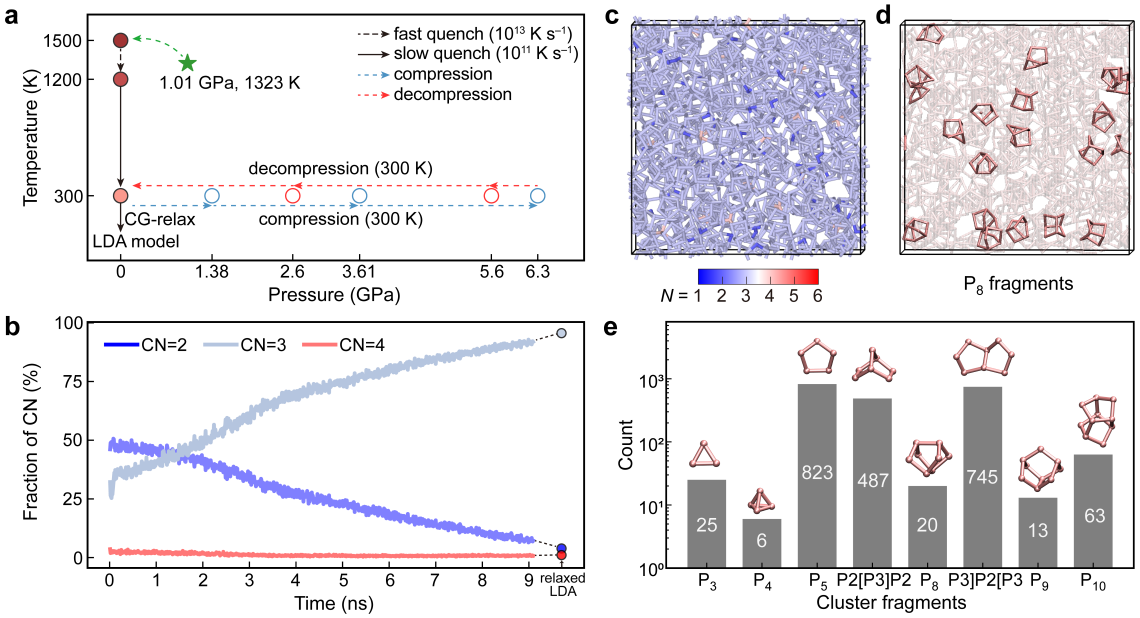


Figure 3.9: The structure of amorphous phosphorus as described by machine-learning-driven simulations. (a) An overview of the protocols used in this work to create the low-density amorphous (LDA) and high-density structural models. Starting from a disordered network liquid, a fast quench (10^{13} K s^{-1}) was carried out to 1200 K, and then a much slower quench (10^{11} K s^{-1}) from 1200 to 300 K. A conjugate-gradient (CG) relaxation was performed for the amorphous structural model before compression and subsequent decompression at 300 K. (b) Coordination numbers (CN) during the quench, based on the count of nearest neighbours with a bond-length cutoff of 2.4 Å. The relaxed LDA structure has 4% two-fold, 95% three-fold, and 1% four-fold coordinated atoms, as indicated by symbols (dotted lines are guides to the eye). (c) Atomic structure of the LDA model, relaxed after the end of the slow quench down to 300 K. The CN is indicated by colour coding. (d) Highlighted P8 cage-like fragments in the simulated LDA model. (e) Counts of relevant cluster fragments as found in the LDA model. Exemplary images of the respective clusters are shown.

studies of a-Si [236].

Notably, the protocol shown in Fig. 3.9a does not mirror experimental conditions directly – but it allows the simulation to probe a variety of possible fragments, on account of the large disorder in the network liquid, and thereby to find local energetic minima. This way, a candidate structural model was obtained which may be (and needs to be!) validated against experiment. Indeed, finding a suitable way of describing the formation of an amorphous phase is a major challenge in the field – no matter whether the simulation was carried out using ML potentials or more established methods. Using amorphous carbon materials as an example, their formation might be described through quenching as a "proxy" [238], through the

annealing of precursors [239, 240], or through direct simulation of the deposition process [195, 241]; in all these cases, the validation of the simulation results by comparison with experimental data has been of central importance. Following the quench to 300 K, the resulting a-P structure was fully relaxed, and referred to as "low-density amorphous" (LDA) in the following, borrowing a term that is often used in the field of disordered silicon [242].

3.4.4 Structural properties of a-P

Figure 3.9b characterizes the changes in coordination number (CN) statistics during the quench: there was a gradual decrease in the number of two-coordinated atoms (CN = 2), and a steady increase for CN = 3. The latter is expected for solid phosphorus as a fifth-main-group element and found in the ambient-pressure crystalline allotropes throughout. Very few four-coordinated atoms (CN = 4), which were regarded as defective, over-coordinated environments, exist in the structural model; most of them are eliminated during the quench process. In the final, relaxed structure, about 95% of the atoms have CN = 3 (Fig. 3.9b–c).

To analyse the cluster fragments, which consist of different combinations of primitive rings, the numbers of the rings and more complex cluster fragments were counted with the help of the R.I.N.G.S. software [243]. The code can transform the atomic structure into a graph with nodes for atoms and links for bonds, and therefore, different sizes of the closed shortest-path rings can be counted. The 3-, 4-, and 5-membered rings are the building units of the more complex cluster fragments, and thus the cluster fragments can be counted based on the arrangements of the rings found by the R.I.N.G.S software. The occurrences of the P8 cage-like fragment, which consists of four fused five-membered rings, are highlighted in Fig. 3.9d; a more comprehensive overview is provided in Fig. 3.9e. The dominant building units in the resulting LDA model were primitive 5-membered rings, labelled "P5". Most of these rings were furthermore involved in the formation of multiple more complex clusters, including the seven-atom "P2[P3]P2" fragment (two fused 5-membered

rings with 3 shared atoms), the eight-atom "P3]P2[P3" one (two fused 5-membered rings with 2 shared atoms), and P10 cages. These fragments, and the way that complex fragments were related to the constituent simpler ones, have been discussed in detail by Böcker and Häser [232]. Only 12.9% of the atoms in the simulation cell (256 out of 1,984) take part in the formation of P8 / P9 cages, much fewer than in crystalline Hittorf's (violet) and fibrous phosphorus (81%) [225, 226]. This is in line with earlier experimental work [228] in which it was stated that P8 cages were present in much reduced numbers in a-P. Besides, the LDA structure only contains few P3 rings and P4 tetrahedra, 25 and 6 respectively, at variance with the model discussed in Ref. 229. Hence, the resulting structural model, generated with a potential "machine-learned" from quantum-mechanical data, has features similar to those proposed by Böcker and Häser on the basis of gas-phase computations [232]. The low abundance of P4 tetrahedra was consistent with reported Raman spectra, in which the intensity of the peak representing P4 was much lower than for those indicating the presence of the P7, P8, and P9 fragments [228, 244, 245].

The calculated mass density of a-P structural model is 2.25 g cm^{-3}, consistent with the reported experimental data ranging from 2.14 to 2.34 g cm^{-3} [229, 246]. Notably, this density is lower than that of Hittorf's (violet) phosphorus (2.36 g cm^{-3}) [225], as well as that of orthorhombic black phosphorus (2.71 g cm^{-3}) [247], indicating a more open network – consistent with the appearance of the structure (Fig. 3.9c–d), which contains covalently linked cluster fragments that are separated from further clusters by abundant voids. The GAP-computed energy of the resulting structural model is 7 kJ mol^{-1} (71 meV $at.^{-1}$) higher, i.e., less favourable, than that of layered black phosphorus, but 8 kJ mol^{-1} (84 meV $at.^{-1}$) lower than that of molecular white phosphorus, and slightly higher than previous predictions for more ordered "nanotubular" allotropes which have been experimentally characterised [248] and studied using dispersion-corrected DFT [249]. The computed energy of the structural model is therefore in line with what one would expect for a

metastable, but experimentally readily accessible amorphous material.

3.4.5 Structural evolution of amorphous phosphorus under pressure

The structure factor of the melt-quenched structural model of a-P was compared to the experimental evidence taken from various previous studies (Fig. 3.10). Despite small deviations, the full ambient-pressure structure factor $S(q)$ computed using ML-driven MD simulations shows fair agreement with the experimental data throughout (Fig. 3.10). Beyond the ambient structure of a-P, pressure-induced structural transitions in the material are of interest and have been characterised using experimental techniques [229, 250, 251]. For example, high-resolution transmission electron microscopy (HRTEM) indicated that nanocrystalline nucleation sites formed at the early stage of compression, which are expected to be mainly driven by the considerable volume decrease [252]. Indeed, equation-of-state measurements also suggested that a-P underwent a substantial volume decrease by 30% when compressed to 8 GPa: a sharp drop in volume was observed at 7.5 GPa [245]. Further (in situ) XRD and Raman spectroscopy experiments showed that, after nucleation under pressure, a-P crystallises into black phosphorus at 7.0 to 8.0 GPa [245, 252].

To study the pressure-dependent behaviour of the cluster fragments on the atomic scale, a compression and subsequent decompression simulations were performed at a temperature of 300 K using the fully relaxed LDA structure. This simulation protocol (cf. Fig. 3.9a) followed the experimental report in Ref. 229, albeit, necessarily, at a faster rate of compression. The LDA structure was first compressed from ambient pressure up to 6.3 GPa, with a rate of 0.02 GPa ps^{-1}; the external pressure was then gradually released at the same rate until it reverted to ambient conditions. The compressed phase at 6.3 GPa was called as the high density amorphous (HDA) phase, and the decompressed phase at ambient pressure was called as "recovered HDA" (rHDA), to emphasise that it was different from LDA phosphorus as generated from the melt-quench process mentioned above –

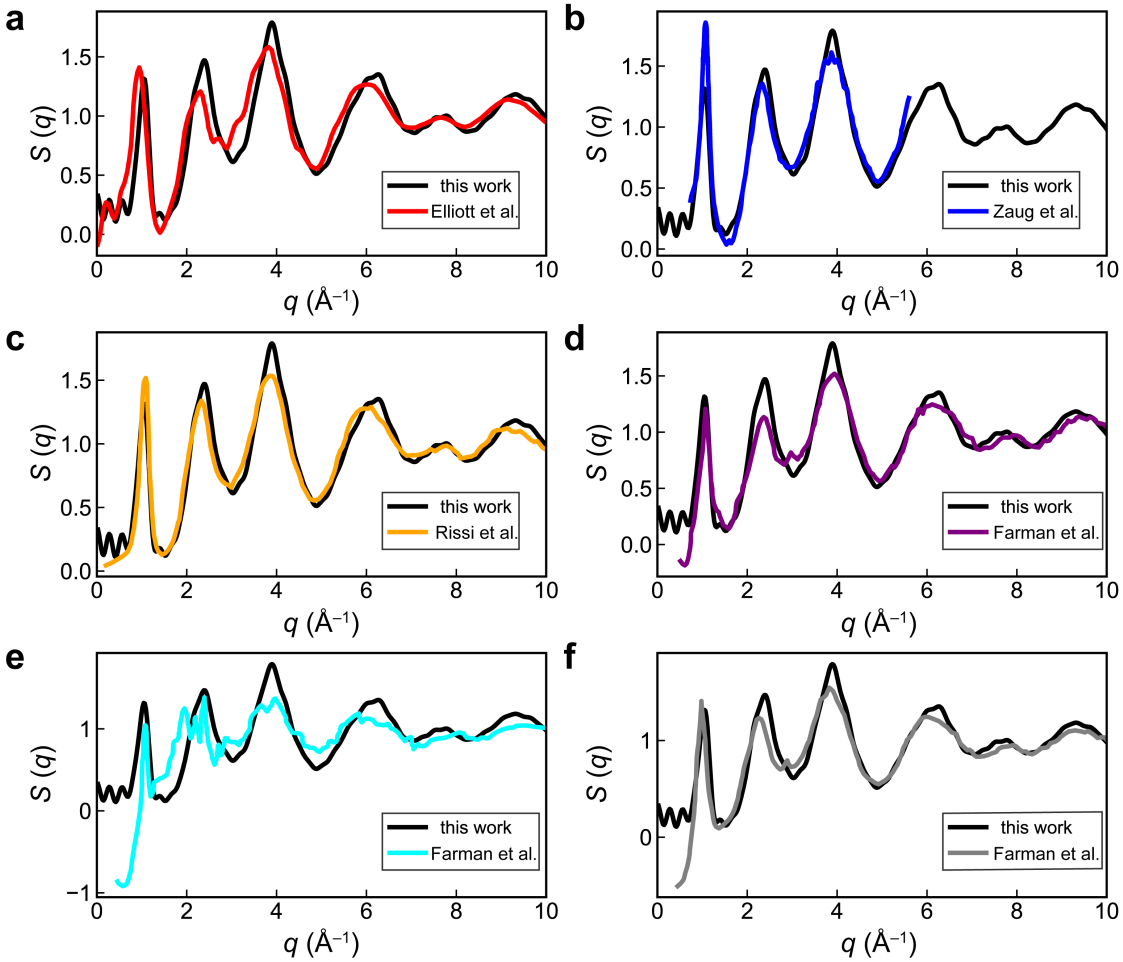


Figure 3.10: The computed ambient-pressure structure factor $S(q)$ for the structural model described in the present work (given in each panel as a black line), as compared to the experimental results given in (a) Elliott et al. (Ref. 224), (b) Zaug at al. (Ref. 229, at 0.01 GPa), (c) Rissi et al. (Ref. 245, at 0.2 GPa), and (d–f) Farman et al. (Ref. 253, where three different samples were investigated). An important point emphasised by this figure is that a-P samples will depend on the synthesis protocol, and their structure (as reflected in the measured structure factor) varies from case to case.

consistent with the high-pressure experimental study of Ref. 229. The results of the compression–decompression simulations were characterised in Fig. 3.11.

The first sharp diffraction peak (FSDP) in the structure factor, $S(q)$, may be affected by multiple aspects of the structure [254, 255], including the distribution of voids as well as the MRO. The FSDP has also played an important role in understanding a-P [229, 245]. The results of $S(q)$ were calculated for the configurations at the same pressure values as reported for the experimental compres-

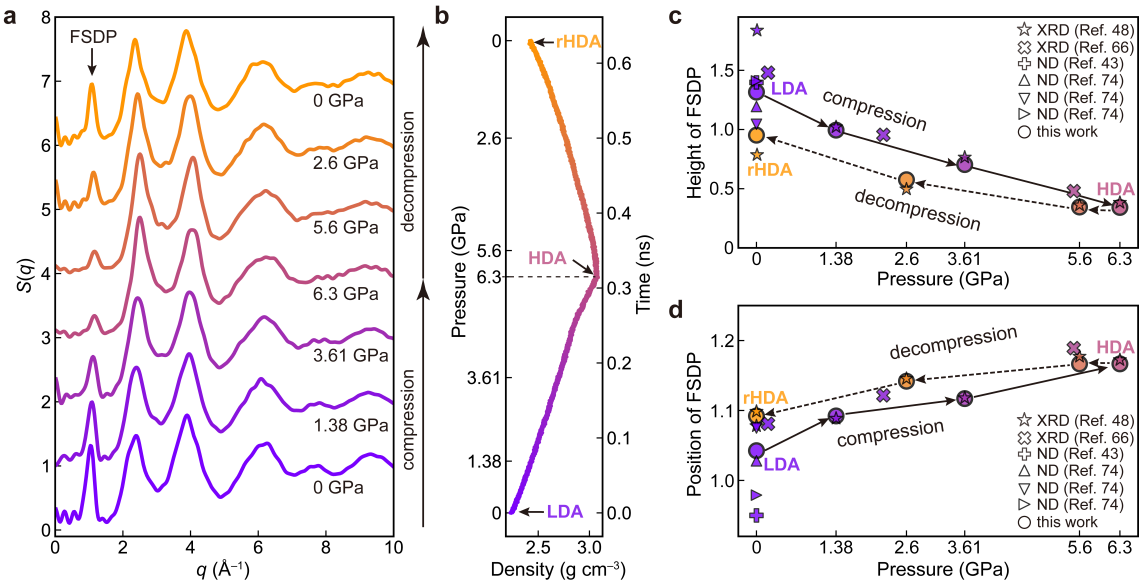


Figure 3.11: Pressure-dependent changes in the structure factor of amorphous phosphorus. (a) Calculated static structure factor, $S(q)$, at ambient temperature during the compression and decompression simulations. The structure factor was calculated based on the Fourier transformation of the computed radial distribution function, g(r), with a real-space truncation distance of 25 Å. Data were vertically offset in steps of +1 for clarity. (b) Calculated mass density during compression and decompression. (c–d) Comparison of the heights and the positions of the FSDP between the calculated structure factors (Fig. 3.11a) and experimental data taken from previous work [229, 229, 245, 253]. Results from different experiments, i.e., X-ray diffraction (XRD) and neutron diffraction (ND) derived FSDP heights and positions, are shown by symbols.

sion–decompression cycle of Ref. 229 (Fig. 3.11a), and additionally the density change under pressure was evaluated (Fig. 3.11b). The ML-driven simulations yield good agreement with earlier experiments in terms of the pressure-dependence of the FSDP:[229, 245] both the simulated intensities (heights) and the positions of the FSDP signals (Fig. 3.11c and 3.11d, respectively) agree well with X-ray and neutron diffraction data. The largest discrepancy in this regard appears for the LDA model compared to ambient-pressure measurements, but these themselves show considerable variation (Fig. 3.10) [224, 229, 245, 253], presumably due in parts to structural differences in various commercial samples. Indeed, a comparative neutron-diffraction study of two a-P samples synthesized at different temperatures revealed differences in the measured $S(q)$, which were interpreted in terms of changes in the intermediate-range order [256]. Besides, it was suggested that

changes in the FSDP can be attributed to the average void size, void spacing, and void density [229], which may differ between the simulated LDA structure and some experimental ones.

During the compression and decompression cycle, a hysteretic recovery was observed, in which the changes in the FSDP only partly reversed upon decompression (Fig. 3.11) – leaving the atomic structure in a denser configuration (Fig. 3.11b). The mass density of the recovered rHDA model after decompression is 2.42 g cm^{-3}, slightly higher than that of Hittorf's phosphorus (2.36 g cm^{-3}; Ref. 225) but still lower than that of layered black phosphorus (2.71 g cm^{-3}; Ref. 247). The right-shifted and less intense FSDP during decompression suggests a more compact and less ordered amorphous structure as compared to the initial one (LDA) undergoing compression, consistent with the density changes shown in Fig. 3.11b and with those reported by Zaug et al. [229]. Note that the sluggish recovery of structural features was also indirectly observed in Raman spectra for compression / decompression experiments [245].

To better understand the pressure-induced behaviour of the cluster fragments, the counts of clusters and higher-membered rings were analysed, as shown in Fig. 3.12. At the start of the compression, up to about 5 GPa, the numbers of these cluster fragments fluctuated slightly but did not change markedly (Fig. 3.12a), and almost no large rings with 7–10 members were found (Fig. 3.12b), suggesting that modest pressure cannot immediately break the local structures. Upon further compression beyond 5 GPa, the number of cluster fragments decreased, which continued up to 6.3 GPa; meanwhile, high-membered rings started to form, which was attributed to the opening of cluster fragments. These observations, again, agreed well with experiments: earlier in situ Raman experiments indicated that pressure up to 3 GPa did not strongly affect the local bonds, whereas all clear peaks found in the ambient Raman spectra became rather weak above 5 GPa, hinting at a significant deformation of the structural building units at higher pressure [244]. In

addition, a very small peak at ≈ 600 cm^{-1} was observed in the Raman spectrum of a-P [244, 245], interpreted as the dominant Raman-active mode of P4 molecules in white P, and this peak disappeared under pressure [244]. It was also found in the simulations that the count of remaining P3 rings and P4 tetrahedra started to decrease at ≈ 4.5 GPa and then quickly disappeared as the simulation reached 6.3 GPa of external pressure (Fig. 3.13a).

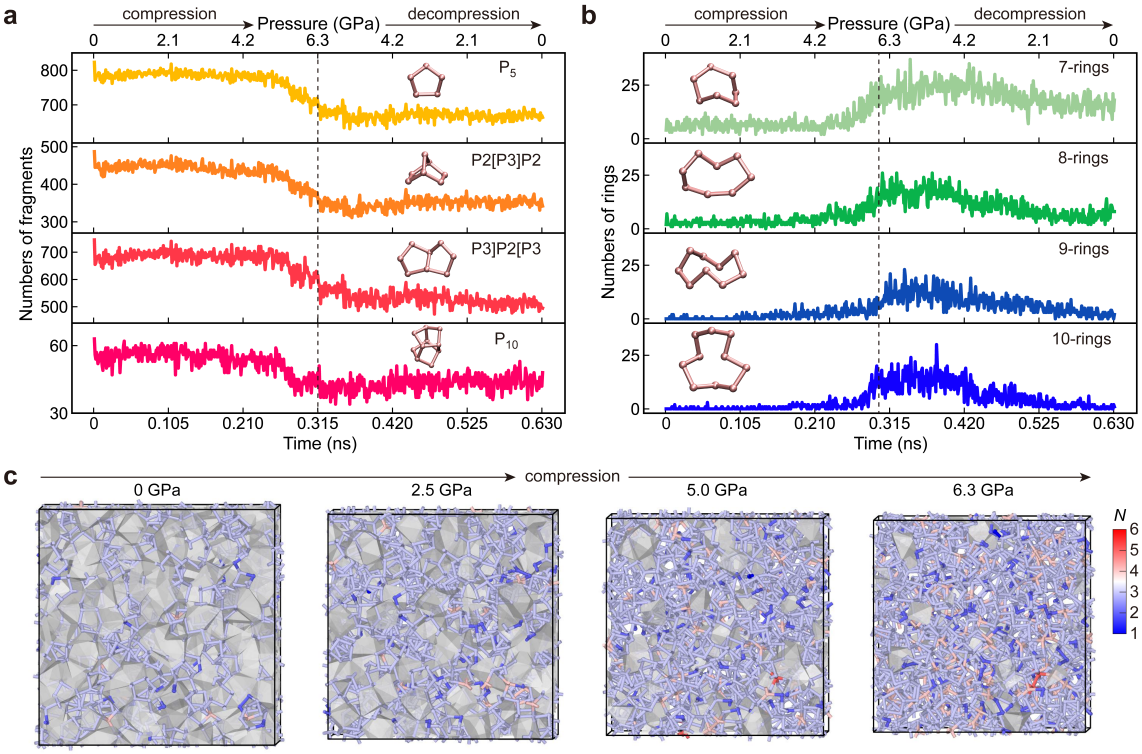


Figure 3.12: The evolution of atomic structure in amorphous phosphorus during simulated compression and decompression. (a) Count of selected cluster fragments, including P5, P2[P3]P2, P3]P2[P3, and P10, and (b) count of large rings, containing 7 to 10 atoms. Vertical dashed lines indicate the change from compression to decompression. Illustrative examples of clusters and rings are shown as insets. (c) Void analysis based on the alpha-shape method [257]. The radius of the probe sphere used here is 2.6 Å, slightly larger than the average nearest-neighbour distance in a-P (cf. Fig. 3.13b). The surface of the void region is rendered in gray.

In the subsequent decompression, from 6.3 GPa back to ambient conditions, those dominant cluster fragments did not immediately become more abundant but gradually re-appeared, further supporting the proposed distinction between LDA and rHDA forms of phosphorus (based on a partial retention of structural features

as pressure is released). Meanwhile, distorted higher-membered rings were barely found at the start of the compression. The count of 7-membered rings increased sharply at about 4.5 GPa, whereas a similar increase takes place in 8-, 9- and 10-membered rings at $\approx$ 5 GPa (Fig. 3.12b). In addition, these larger rings slowly disappeared when pressure is gradually released, and many of them remained after decompression and in the rHDA phase, in contrast to the LDA phosphorus.

A void analysis was performed to better understand the evolution of the MRO in a-P, and the results are shown in Fig. 3.12c. The void regions in the computed structures (indicated by gray surfaces) are identified based on the alpha-shape method [257] as implemented in the Ovito software [258]. This method is based on the construction of a surface separating filled spatial regions and empty regions, assuming that all atoms are points (i.e., without a radius). A probe radius of 2.6 Å was used to detect the voids. Upon compression to 5.0 GPa, it was found that that most voids in LDA phosphorus were gradually squeezed out, and different cluster fragments were brought closer to each other yet remained intact, leading to the formation of relatively long P$\cdots$P contacts (2.4–2.8 Å) between neighbouring clusters. This was further evidenced by the gradually increased magnitude of the radial distribution function around its first radial minimum during compression (Fig. 3.13b). As pressure further increased (from about 5.0 to 6.3 GPa), the clusters started to open, leading to the disappearance of LDA-like fragments (Fig. 3.12a) and a simultaneous increase in the count of larger rings (Fig. 3.12b).

3.4.6 Creating small-scale models of a-P

Just like there are open questions about many crystalline phosphorus allotropes [249], there remain fundamental questions about a-P. For example, how does the degree of local structural ordering determine the energetic (meta-) stability? And where does a-P fall within the previously introduced first-principles stability range of the crystalline phosphorus allotropes [249]?

Those questions will be precisely addressed by using a combination of ML-driven

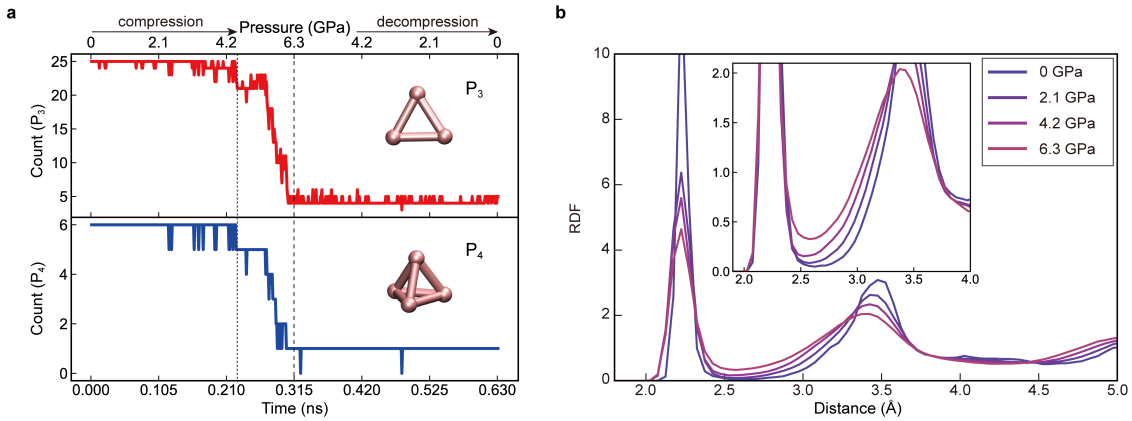


Figure 3.13: Structural analysis during the compression process. (a) The evolution of the count of P3 rings and P4 pyramids during simulated compression of the LDA phosphorus structure. The dashed line in the middle represents the change from compression to decompression. Neither of these fragments is highly abundant in the initial structure, and a sharp decrease in the count of both P3 and P4 units occurs at about 4.5 GPa, as indicated by the dotted line. (b) Calculated radial distribution functions (RDF) of a-P during the compression from 0 GPa to 6.3 GPa. The region including the first minimum is enlarged in the inset. Each RDF result is averaged over a 5-ps trajectory (50 successive snapshots with a time-interval of 0.1 ps) at the given pressure point during the compression.

MD simulations and first-principles DFT including many-body-dispersion (MBD) corrections. Three representative structural models were introduced for a-P – large enough to describe the complex structures, yet small enough to enable full first-principles investigations. New insight was derived into the role of local structural order in the energetic stability of a-P, the nature of defects, and the balance between covalent and vdW dispersion interactions in phosphorus allotropes.

The melt-quench MD simulations were carried out, using the GAP potential for P [178] to create structural models of a-P [211]. The simulations started from a metastable liquid, rather than the fluid (consisting of P4 molecules) which was stable at low pressure – this was done on purpose, because the large structural diversity in the metastable liquid allowed the simulation to rapidly explore relevant local structural fragments. The protocol followed the same melt-quench process mentioned in the previous section and in the earlier work [211], but the new simulations were now carried out in smaller simulation cells, with the aim to enable subsequent DFT studies, and in three parallel runs.

These structural models each contained 248 atoms with mass densities of 2.5, 2.5, and 2.4 g cm^{-3}, respectively; their structures represented different stages of the liquid-to-liquid phase transition [251, 259]. Both P molecules and covalently connected networks were observed in the former two models, and the third almost fully resembles the network liquid without any P tetrahedra. The models were rapidly de-compressed from $\approx$ 1 to 0 GPa and annealed at 1,500 K over 10 ps, and then quenched to 1,200 K at a rate of 10^{13} K s^{-1}. After that, the structural models were slowly cooled down from 1,200 K to 100 K at a rate of 10^{11} K s^{-1}, thereby forming amorphous networks. During quenching, the number of fully three-coordinated atoms steadily increased. It has been previously shown in the previous large-scale melt-quench simulation (in the previous section) that this process allowed the simulation to explore the relevant configurational space, leading to a structural model consistent with previous experimental data (including structure factors at ambient and high pressure; see Ref. 211 and the previous discussions). The a-P models were further relaxed successively using GAP and DFT (PBE+D3).

Having three small-scale models of a-P allowed to take advantage of the stochastic nature of the process: with the same simulation protocol but different starting configurations, these structures over millions of simulation steps evolved into rather distinct representative disordered networks, which were labelled as **1** to **3** (Fig. 3.14).

Those a-P models, after DFT-based structural optimisation, contained mostly three-fold coordinated atoms, in line with standard valence rules; there were only $\approx$ 1% of over-coordinated ($N = 4$) and $\approx$ 2% of under-coordinated atoms ($N = 2$). There was a different degree of structural ordering in the three a-P models (Fig. 3.14a): **1** showed a more disordered, random network, whereas the arrangement of fragments in **3** followed patterns resembling those in violet and fibrous phosphorus (viz. tubular networks). Hence, the three models can be viewed as suitable representatives for the range of structural variety that one might expect in an experimental sample. Similar to the previously simulated large-scale a-P model [211], the three

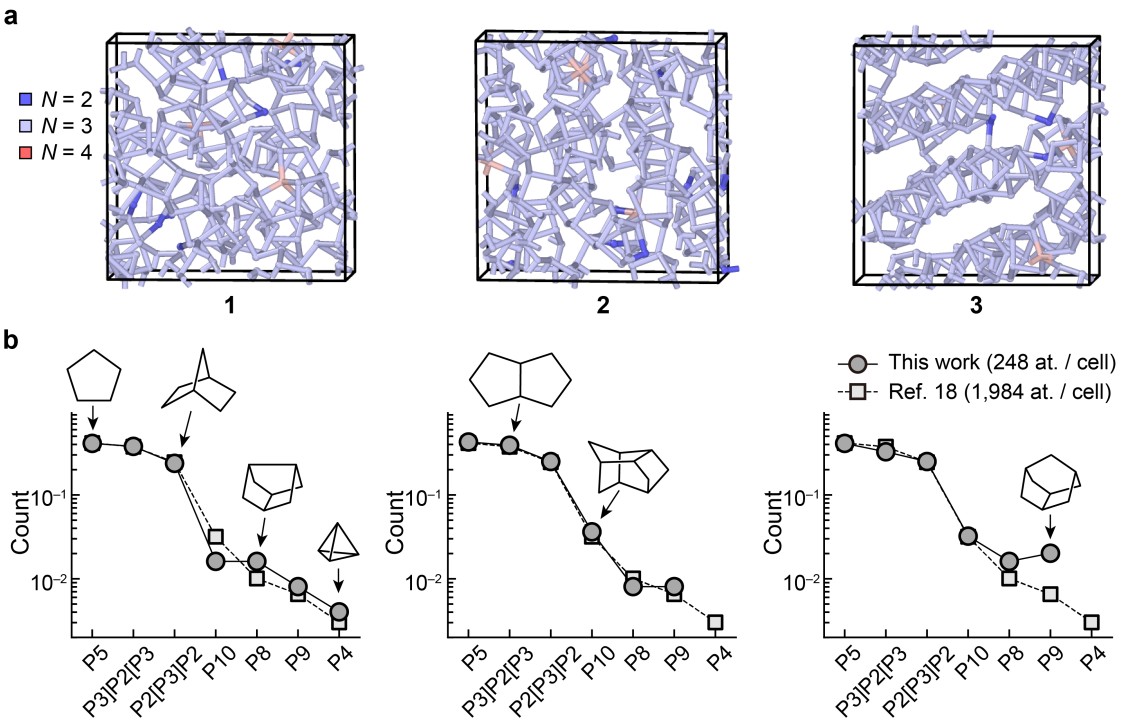


Figure 3.14: Structural models of amorphous phosphorus. (a) Structures of a-P as obtained from ML-driven simulations and subsequent DFT structural optimisation. Three independent structural models were generated containing 248 atoms each, using the same simulation protocol in each case. The structures are labelled **1** to **3**, with an increasing degree of structural order. Colour-coding indicates the coordination numbers, viz. $N = 2$ (blue), $N = 3$ (pale blue), and $N = 4$ (pink), determined using a 2.4 Å cut-off. (b) Relative abundance (per-atom count) of cluster fragments in the a-P models (circles), as compared to the result of a similar but larger-scale 1,984-atom simulation in Ref. 211 (squares). Solid and dashed lines are guides to the eye. Sketches of selected fragments are shown, following Ref. 232

models created here contained abundant five-membered rings, and more complex cluster fragments that are made up of those (e.g., the P8 cage found in violet and fibrous P consists of four fused five-membered rings). The same notation was used as in Section 3.4.2 and Ref. 232 to label those fragments.

The distribution of cluster fragments (Fig. 3.14b) confirmed that three new small-scale models were overall consistent with the previously validated, larger-scale model. A comparatively higher count of P8 and P9 cages in **3** implies a greater similarity with the closely related crystalline forms, violet and fibrous phosphorus. Such cage-like fragments have also been used as building blocks of predicted phosphorus allotropes with more complex local environments, via a random structure

search [260]. No P4 units were observed in **2** and **3**, implying pure network, rather than molecular, structures. The overall agreement between the present structures and the one from Ref. 211 underlined that purpose-tailored simulations were possible with ML potentials: these enable ultra-large-scale simulations, but also the use of smaller-scale models which were amenable to subsequent first-principles DFT analyses of energetics and chemical bonding.

3.4.7 Energetic stability of phosphorus allotropes

To compare the energetic stability of different phosphorus phases, including the a-P structures (**1** to **3**) and relevant crystalline forms, the energies of fully DFT-optimised structures were computed for those modifications (Fig. 3.15), using projector augmented-wave (PAW) potentials as implemented in the Vienna Ab initio Simulation Package (VASP) [261–263]. Corrections for vdW interactions were needed to accurately describe longer-range interactions in phosphorus modifications [178, 264]: here the "D3" method was used [265, 266] for structural optimisations, and the many-body dispersion (MBD) method [267, 268] was implemented for subsequent single-point computations. Both methods can describe the exfoliation of black phosphorus remarkably well [269], and MBD also captured the unit-cell volume of black phosphorus in almost quantitative agreement with experiment, and the exfoliation energy as compared to higher-level quantum-chemistry methods [178], viz. quantum Monte Carlo [270] and coupled-cluster theory [271].

The computed energies provided a full stability range of phosphorus allotropes. Among the phases studied, white phosphorus is the least energetically stable (Fig. 3.15). Violet, fibrous, and black phosphorus are all predicted to have close energies, with differences less than 0.3 kJ mol^{-1}, although the structure of black phosphorus differs from that of the other two. In fact, such energetic near-degeneracy is not only seen with HSE06+MBD, but was reported using various DFT methods [235, 249]. Following Ref. 249, three models of P nanorods were also generated (referred to as $n1$, $n2$, and $n3$), by removing the copper and iodine atoms from $(CuI)_8P_{12}$, $(CuI)_3P_{12}$,

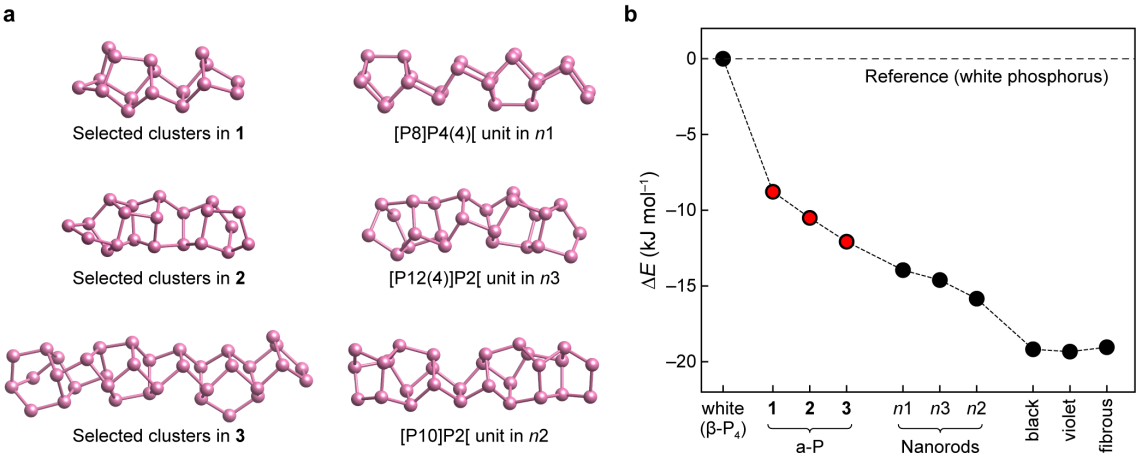


Figure 3.15: The energetic stability range of phosphorus modifications, including the well-known white, black, violet, and fibrous forms, as well as structural models of Pfitzner's nanorods (*n*1 to *n*3) based on Refs. [248] and [249], and the new a-P structural models **1** to **3** generated in the present work. (a) Selected, local structural fragments in the a-P models compared to the building units of nanorods. (b) Computed energies, given relative to the standard state, viz. white phosphorus, and based on DFT computations (HSE06+MBD).

and $(CuI)_2P_{14}$, respectively, and then relaxing the structures using DFT. The Vienna Ab initio Simulation Package (VASP) [263, 272] was used together with projector augmented-wave (PAW) [261, 262] pseudopotentials. The PBE+D3 method [265, 266, 273] was used for structural relaxations. The plane-wave energy cut-off was 500 eV, the energy tolerance for SCF convergence was 10^{-7} eV per cell, and the force tolerance for structural relaxation was 0.01 eV Å^{-1}. Gaussian smearing with a width of 0.05 eV was used to determine partial occupancies during SCF cycles. In structural relaxation, all lattice parameters and atomic coordinates were optimised. A k-point grid with a largest allowed spacing of 0.1 Å^{-1} along each reciprocal lattice vector was used for structural relaxation for crystalline modifications, whereas for a-P models, only the Γ-point was used (corresponding to a "1 1 1" grid).

Three a-P models (**1** to **3**), energetically, sit in between white P and the other crystalline forms, and with increased structural ordering (Fig. 3.14a), the energies of these three models decrease. Table 3.1 shows that the computed mass densities of different allotropes agree well with experimental data, and the computed energy ranking of the crystalline forms is consistent with previous studies [249, 264] – the

combined ML- and DFT-based simulation approach has now allowed to extend this ranking to the previously poorly understood amorphous form.

Bachhuber et al. have previously studied dispersion interactions in phosphorus allotropes, using pairwise dispersion-correction methods available at the time [275]. With the consideration of advanced MBD corrections, the vdW contributions to the overall energy also vary across the different phosphorus modifications (Table 3.1): white phosphorus shows the smallest absolute vdW contribution (lowering the total energy by about 14 kJ mol^{-1}), whereas black phosphorus has the largest value ($\approx$ 20 kJ mol^{-1}). All structures studied that contained cage-like motifs have a similar level of vdW contributions, of about 16–17 kJ mol^{-1} in stabilisation relative to the pure, uncorrected HSE06-DFT energy. These results can be understood from the different interspaces of building fragments in molecular and network solids: in white phosphorus, the P4 tetrahedra are well-separated from each other, and the distance between building units is longer than in other modifications, resulting in smaller dispersion contributions. This result is intuitive in the sense that vdW forces approximately decay with the sixth power of the distance, but it is still notable that

Table 3.1: Properties of the a-P models **1** to **3** and comparison to crystalline phases. The standard phase, white phosphorus, was set as the reference. The energies of other structures, computed using HSE06+MBD, are given relative to white phosphorus. The van der Waals (vdW) contribution was listed separately to the computed total energy for each structure.

	Energetics (kJ mol^{-1})		ρ (g cm^{-3})	
	ΔE (HSE06+MBD)	vdW (MBD) contribution	DFT	Expt.
White (β-P_4)	$\pm$ 0 (reference)	-13.7	1.89	1.98 [274]
1	-8.8	-16.4	2.25	2.14 to
2	-10.5	-16.5	2.27	2.34
3	-12.1	-16.5	2.29	[229, 246]
$n1$	-14.0	-15.9	2.18	
$n2$	-15.8	-16.8	2.29	
$n3$	-14.6	-16.4	2.23	
violet	-19.3	-17.1	2.32	2.36 [225]
fibrous	-19.0	-17.2	2.31	2.37 [226]
black	-19.2	-19.9	2.64	2.69 [247]

the "molecular" solid, P4, is less strongly vdW bonded than any of the black, violet, fibrous, or amorphous forms.

In short, this section has shown how computer simulations using an ML interatomic potential have provided an atomic-scale picture of the local structure and the pressure-dependent structural changes in amorphous phosphorus. On the one hand, the agreement between simulated and (previously) experimentally obtained structure factors provides a firm basis for interpreting the simulation result, and it enables the study of atomic-scale structural changes that are not directly accessible to experimental observation. On the other hand, the ML-driven simulations provide a starting point for future work on the functional material system, i.e., the electrochemical insertion of alkali-metal ions in phosphorus-based battery anodes.

Looking beyond the new findings regarding phosphorus, the ML-based results shown in this chapter provide an example of the growing importance of ML-driven modelling for understanding structural changes in disordered functional materials. Such mechanisms are important both for fundamental insight and for practical applications in devices, and they are poised to become much better understood with the large-scale and quantum-mechanically accurate simulations that ML methods are increasingly making possible.

3.5 Quantification of per-atom crystallinity

Studies of crystallisation in phase-change materials are based on the quantification of per-atom crystallinity, which characterises the whole crystallisation (nucleation and growth) process. Here, an example was given to illustrate how quantifying per-atom crystallinity can help the understanding of crystallisation dynamics in GST alloys. Part of this section is based on a previously published work in Adv. Mater. (Ref. 114), where I am the co-first author. Some texts and figures have been reused. I would like to acknowledge the useful discussions with Yazhi Xu, XuDong Wang, Wei Zhang, En Ma, Volker L. Deringer, and Riccardo Mazzarello.

To start with, two 1008-atom models of rock-salt-type $Ge_1Sb_2Te_4$ were created

in an orthorhombic supercell of size 2.954×2.558×4.178 nm^3 (144 Ge, 288 Sb and 576 Te atoms, and 144 vacant sites), corresponding to the experimental crystalline density ($a = 6.03$ Å in the conventional cubic unit cell). The density was fixed during the whole crystallisation process. This scheme is relevant to phase-change random access memory (PCRAM) devices, wherein the crystalline matrix surrounding the amorphous region constrains the volume available during crystallisation.

This model size was shown to be large enough to capture strong localisation of electrons in Ref. 11. With two (111) crystalline layers being fixed, the system was first randomised above 2000 K for 10 ps, and then quenched to 1100 K in 6 ps. It was held at 1100 K for 30 ps and then quenched to 300 K with a cooling rate of 10 K ps^{-1} to create a planar amorphous–crystalline interface. Subsequently, the models were re-heated and kept at ≈630 K for about 600 ps, during which they both crystallised into a cubic rock-salt phase. The two models are denoted as "model 1" and "model 2" in the following. Some snapshots of the two crystallisation trajectories were shown in Fig. 3.16. In both models, crystallisation is dominated by the growth from the interface and no nucleation event takes place during the simulation. This is because nucleation is a stochastic process, and rapid crystal growth from the interfaces led to complete crystallisation before a sizeable nucleus could form.

In previous studies of crystallisation of GST alloys, the "dot-product" bond order parameter, q^4_{dot}, is one of the most widely used indicators for per-atom crystallinity. The detailed description of the definition and use of q^4_{dot} was discussed in Ref. 109. Using q^4_{dot}, the per-atom crystallinity was calculated for each atom in the structural model. This parameter distinguishes crystal-like and amorphous-like environments. Based on the results shown in Fig. 3.17, an atom is crystal-like if its value of q^4_{dot} is larger than 0.45.

The growth velocity v_g is then determined by computing the total volume of the crystalline region V and the total area S of the interface between the amorphous

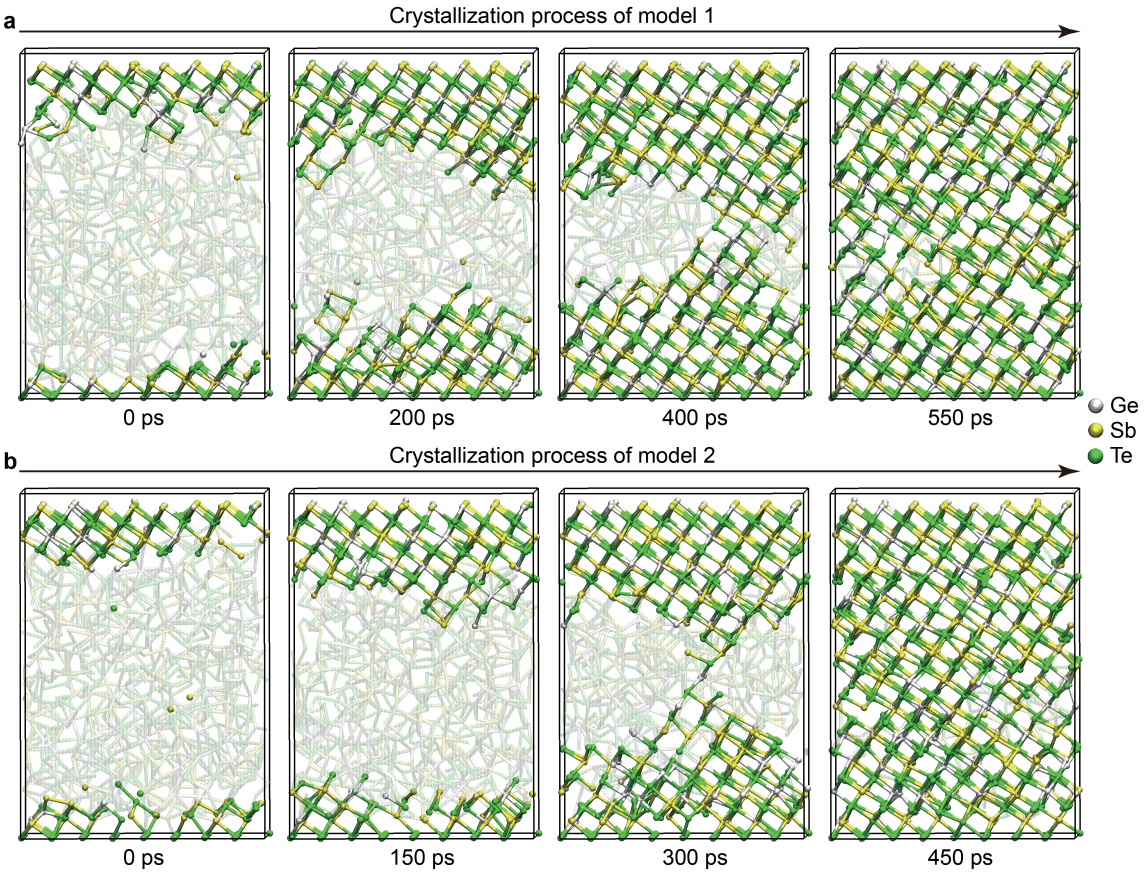


Figure 3.16: *Ab initio* simulated crystallization of a 1,008-atom model of the $Ge_1Sb_2Te_4$ phase-change material. Snapshots of the crystallization of (**a**) model 1 and (**b**) model 2 both contain an amorphous–crystalline interface. Atoms assigned as "crystalline" by the ${q_4}^{\mathrm{dot}}$ analysis are rendered as solid (${q_4}^{\mathrm{dot}} > 0.45$); atoms assigned as "amorphous" are rendered as semi-transparent (${q_4}^{\mathrm{dot}} < 0.45$). The threshold value for this assignment is ${q_4}^{\mathrm{dot}} = 0.45$. Some crystalline protrusions are found on the growth interface upon crystallization (that is, the growth front is not fully smooth).

and the crystalline region at each time step, as also described in the previous work [109]. The v_g as a function of time t can be seen as the temporary variation of crystalline volumes V, divided by the crystal-amorphous interface areas S, shown as follows:

$$v_g(t) = \frac{1}{S} \cdot \frac{dV}{dt}$$

The crystal-like particles can be determined based on the distributions of quantified crystallinity, in which a selection value can clearly separate the crystal-like and amorphous-like particles. Note that some crystal-like (or amorphous-like) atoms are

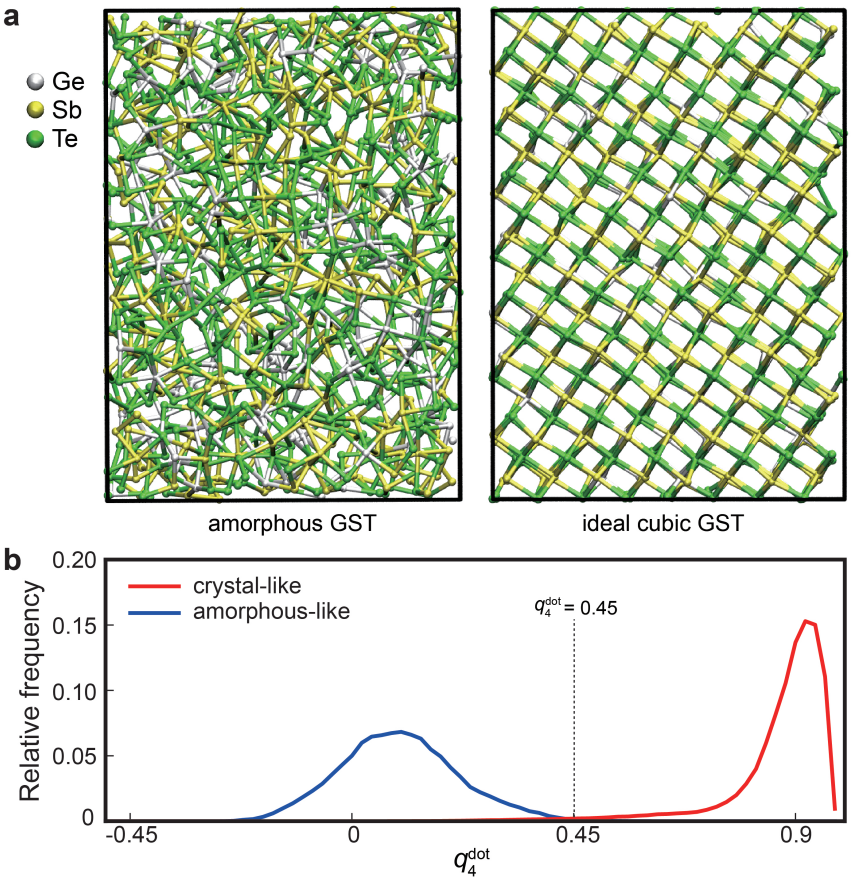


Figure 3.17: (a) Examples of typical crystalline and amorphous structures. Both models have the same atomic density, i.e., the experimental crystalline density, which corresponds to the one of the crystallization models used in this work. The amorphous structure was generated by the melt-quench protocol, as discussed in the main text, without fixing any crystalline layers. (b) The distribution of q^4_{dot} values for the atoms of the crystalline and amorphous models, averaged over 1000 snapshots respectively. The plot indicates that the threshold value $q^4_{dot} = 0.45$ can be used to properly distinguish the crystal-like and amorphous-like particles.

separated from the crystalline (or non-crystalline) matrix. The presence of these atoms, mainly on account of thermal fluctuations, can be seen as a "noise" during the recrystallisation process, leading to an inaccurately estimated crystalline volume and an enlarged growth interface. Hence, as a correction for this "noise", the crystalline volumes V were manually calibrated and the areas of growth interface S, created by these "noise" atoms, were deducted, before calculating v_g. The raw data of volumes V and areas S are smoothed using a Savitzky–Golay filter with a time window of 20 ps for the calculation of growth velocity.

The quantities V and S are obtained by summing up, respectively, the volumes

and the relevant surface areas of the Voronoi polyhedron [276] around the atoms that belong to the crystal (Fig. 3.18). Notably some atoms are occasionally identified as crystal-like in the non-crystalline region, due to thermal fluctuations (Fig. 3.16), which in fact should be considered as "noise" during the crystallisation. Hence, a correction is performed on these atoms when calculating V and S. Averaging over time and two structural models, we obtain a growth velocity of $\approx$1.5 m s^{-1} at 630 K (Fig. 3.18). The narrow growth interfaces and high atomic mobility in the amorphous-like parts (Fig. 3.17) enable frequent atomic impinging and attaching on the crystal-like boundary, leading to fast growth. These observations are consistent with the previous crystallisation simulations on $Ge_2Sb_2Te_5$ and $Ag_4In_3Sb_{67}Te_{26}$ PCMs [1, 113].

The large-scale structural models provide additional details about the crystallisation process that were not clearly captured in previous AIMD simulations of smaller model sizes. In particular, the amorphous-crystal interfaces become notably extended along the growth direction during crystal growth, due to the formation of some crystalline "protrusions", which hints at roughness of the growth front at the nanoscale (Fig. 3.16). Furthermore, both recrystallised models contain a non-negligible amount of antisite defects (Ge / Sb atoms on Te sites, or vice versa). The formation of "wrong bonds" was also observed in previous crystallization simulations of GST with and without an embedded nucleus [103–105]. Unlike local *structural* disorder, these antisite defects, as *chemical* disorder, cannot be directly identified by q^4_{dot}, which only relies on bond lengths and angles.

To cope with this challenge, I have developed a protocol based on the smooth overlap of atomic positions (SOAP) [170] similarity kernel to quantify progressive changes in both the crystallinity (here referred to as "SOAP crystallinity") and the *chemical* disorder (antisite defects) during the crystallization process. Initially used as a descriptor in the fitting of machine learning (ML) based interatomic potentials [170], the SOAP formalism has also been shown to be useful for analysing many

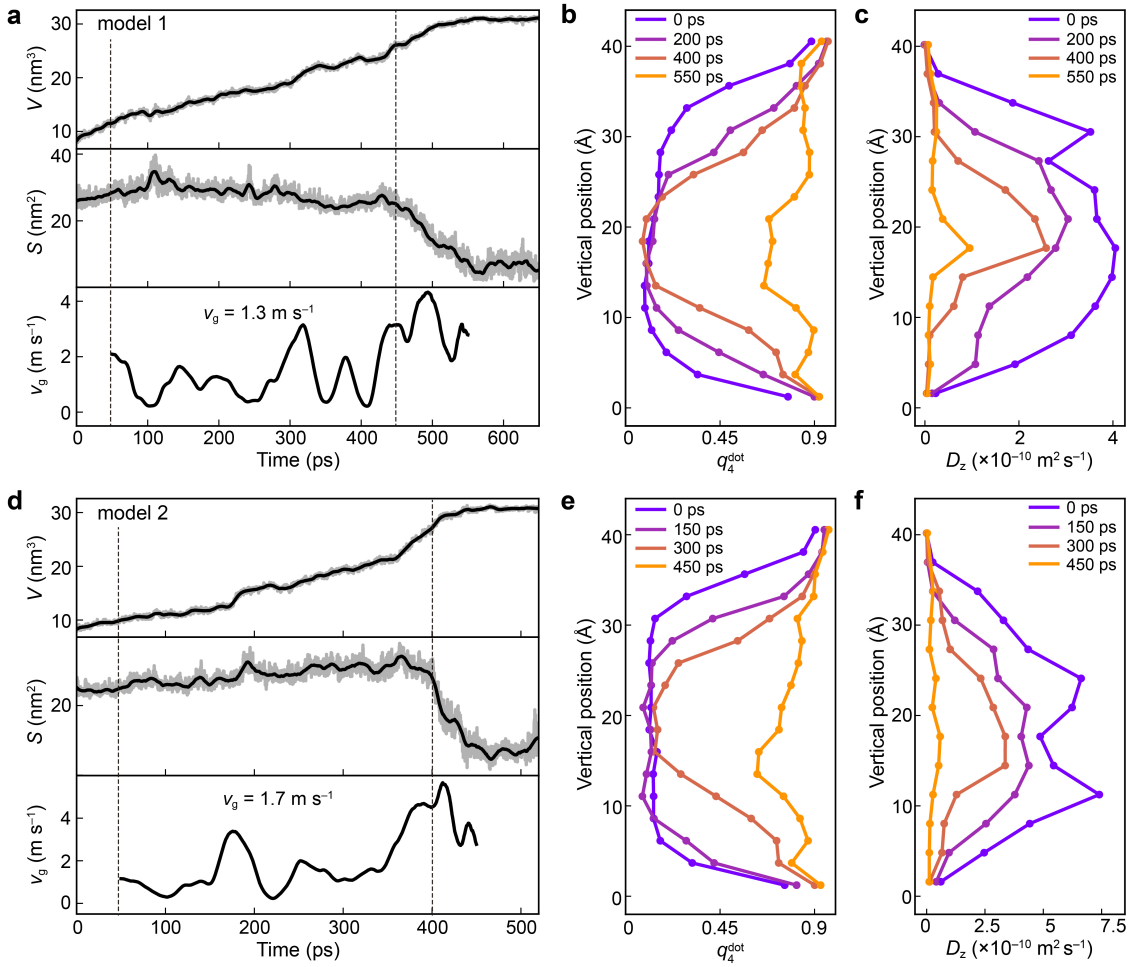


Figure 3.18: Analysis of the crystallization process of model 1 (a-c) and model 2 (d-f). (a) Volume of the crystalline region V, area S of the interface between the amorphous and the crystalline region, and crystal growth velocity v_g as a function of time for model 1. The black solid lines represent a smoothed result over the raw data ($grey$ lines), using a Savitzky–Golay filter with a time window of 20 ps. The time window between two dashed lines in (a) and (d) is used to calculate the growth velocity. The average growth velocity was calculated from the data in the time window (indicated by the dashed lines). (b) Profiles of q^4_{dot} along the direction of growth z for the four structures of model 1 shown in Fig. 3.16a. At the start of the crystallization, the profile of q^4_{dot} shows a sub-nm narrow interface, which extends over approximately 4 atomic layers (7 Å). (c) Profiles of the diffusion coefficient D_z during the crystallization simulation of model 1. The profile is derived from the mean square displacements in the z -direction, which are calculated by averaging over time (10 ps) and over the atoms in each slab. It is evident that, within the narrow interfaces, the atomic diffusivity decays quickly from the amorphous-like to the crystal-like region. (d-f) The same quantities are also calculated for model 2, and similar sub-nm narrow growth interfaces are also observed. Averaged over two models, the estimated growth velocity from the AIMD simulations is 1.03 m s^{-1}.

aspects of atomistic structure [169, 277], including the similarity between amorphous and crystalline materials [147, 206, 278, 279].

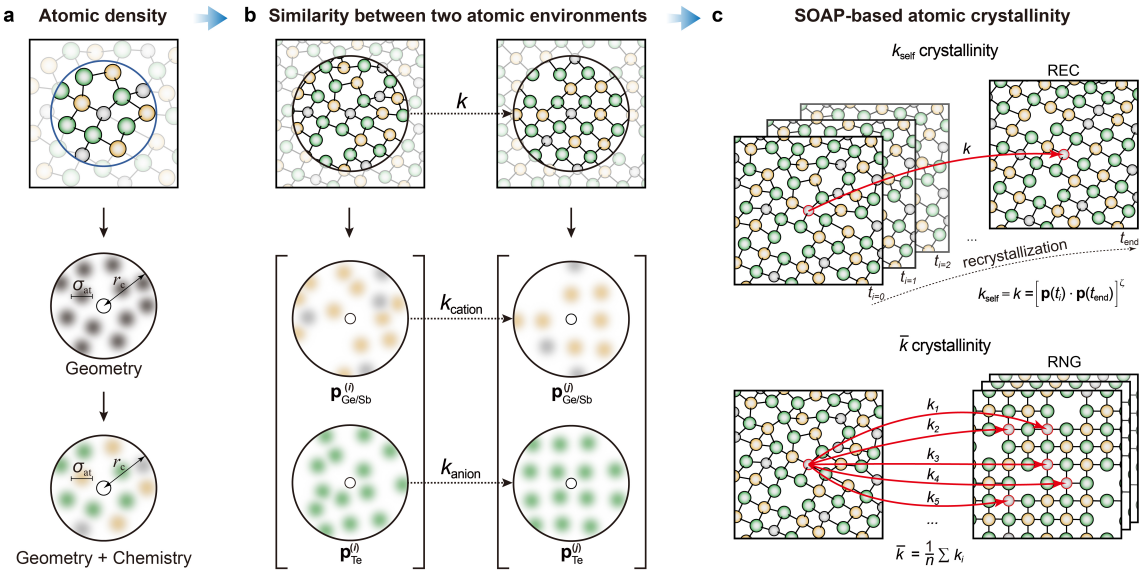


Figure 3.19: SOAP-based characterization of the recrystallization process. (a) Construction of local atomic density for a given atom. A Gaussian function is placed on each atomic position within a sphere of cutoff r_c. The chemical information can be analysed by constructing the element-resolved local atomic density (ELAD). (b) The calculation of similarity between two atomic environments. In this work, Ge and Sb (cation) atoms are considered in one ELAD, and Te (anion) atoms in the other one. The similarity k between two environments can be computed via a dot product (indicated by dashed arrows) of the power-spectrum vectors p of these two ELADs, and then averaged over all pairs. (c) The principles of k_{self} and $\overline{k}$ crystallinity measures. k_{self} characterizes the similarity of a given atom in the recrystallized model with respect to the fully recrystallized structure ("REC"); here, a simple (single) dot-product kernel is shown for simplicity. $\overline{k}$ is computed as an average over the similarities of this atom with all individual atoms of the same element in the reference structure (here, "RNG").

To obtain the "fingerprint" of an atomic environment (within a sphere of cutoff r_{c}) for a given atom, the local density is constructed via a sum of Gaussian functions with broadness σ_{at} placed on each neighbour of the central atom (Fig. 3.19a). Both r_{c} and σ_{at} need to be carefully calibrated (Fig. 3.20). The chemical order within this atomic environment is of particular interest, and densities of different species give rise to what was called the "element-resolved local atomic density" (ELAD), in (nominal) analogy to the element-resolved atomic structures of $Ge_2Sb_2Te_5$ presented in Ref. 10, and the element-resolved local lattice distortion discussed in Ref. 280. The similarity k between two atomic environments can be computed by matching separately these n ELADs of different species, and then averaging (Fig. 3.19b). We note that two or more atomic species can be taken into account in one ELAD, if these

species are chemically equivalent to each other – as we do here for Ge and Sb, which both occupy cation-like sites largely interchangeably with one another, whereas Te is defined as the other ELAD because the swap between Te and cation-like atoms leads to antisite defects (Fig. 3.19b).

To now define a "SOAP crystallinity" for atoms during crystallization, a straightforward way is to calculate the SOAP similarity between the current state and the final state of each individual atom. This crystallinity measure, which was called k_{self}, is based on the fully recrystallized structure from the same AIMD simulation (referred to as the "REC" model; Fig. 3.20a). Another, more general way of defining a "SOAP crystallinity" is to calculate the average similarity of a given atom with respect to all atoms of the same species in a selected ensemble of reference structures. This quantity, which was called $\overline{k}$, requires the choice of a separate reference (Fig. 3.19c). Given that both models crystallized into a single cubic phase and that the AIMD-simulated REC models contain local distortions and antisite defects, we here use the $\overline{k}$ crystallinity and consider the reference structures to be three idealized rock-salt-like $Ge_1Sb_2Te_4$ models with randomly distributed Ge/Sb/vacancies, independently constructed with the help of a pseudo-random number generator (referred to as "RNG" model; Fig. 3.20b).

More specifically, to quantify both the chemical and geometric similarity of two atomic environments (ξ^A and ξ^A) up to a cutoff r_c, the SOAP approach [170] is used to construct suitable kernels. To address chemical complexity, the element-resolved local atomic density (ELAD) was considered which includes different species (Ge/Sb or Te) separately in the local atomic environment (Fig. 3.19a), noting the similarity to "alchemical" SOAP-based kernels in Ref. 277. A Gaussian function with broadness σ_{at} is placed on each atomic position within the cutoff, and the neighbour density is then expanded in a basis of radial functions $R_n(r)$ and spherical harmonics $Y_l^m(r)$, as shown in Equation. 3.11. The convergence parameters $n \leq 16$ and $l \leq 16$ were used here. The similarity between two atomic environments is obtained by

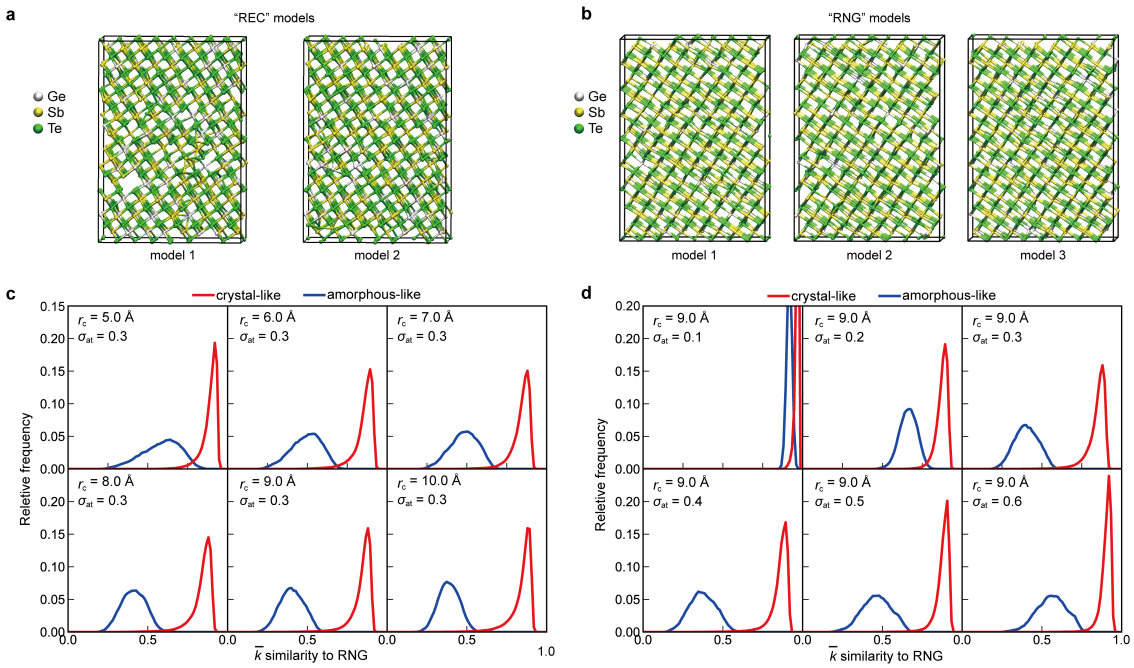


Figure 3.20: Two reference structures mentioned in the main text and the role of the SOAP (hyper-) parameters in the analysis of crystal-like and amorphous-like regions in simulations of GST. (a) "REC": fully recrystallized cubic GST model 1 and model 2, which were quenched from the high temperatures and relaxed at 0 K, and (b) "RNG": idealized cubic GST with randomly distributed Ge/Sb/vacancies, which was created with the help of a random number generator.(c) Tests for the cutoff r_c. The σ_{at} value was fixed to be 0.3 and six different r_c values were employed, ranging from 5.0 Å to 10.0 Å. It is found that $r_c > 8.0$ can well distinguish the amorphous-like and crystal-like particles. (d) Tests for σ_{at}. The r_c value was fixed to be 6.2 Å and six different σ_{at} values were tested, ranging from 0.1 to 0.6. The amorphous-like and crystal-like particles can be well separated when using $\sigma_{\text{at}} = 0.3$ or above. Given that larger values r_c of σ_{at} lead to more expensive computations, the values of $r_c = 9.0$ and $\sigma_{\text{at}} = 0.3$ were selected.

calculating the dot product of the power-spectrum vectors $\mathbf{p}_\alpha$ of two ELADs of the same element (or, here, group of elements, viz. Ge/Sb) [277]. An average was then computed over n species:

$$k(\xi^A, \xi^B) = \frac{1}{n} \sum_{\alpha}^{n} \mathbf{p}_\alpha(\xi^A) \cdot \mathbf{p}_\alpha(\xi^B).$$

Given the reference structures to be three RNG models, the SOAP crystallinity $\overline{k}$ of an individual atom x in the recrystallized model was thus calculated with respect to the reference structures, defined as an average over the similarities of this atom to all the N atomic environments of the same species in the reference structure:

$$\bar{k}(\xi^x, \mathrm{RNG}) = \frac{1}{N} \sum_{\xi^i \epsilon \mathrm{RNG}}^{N} k(\xi^x, \xi^i).$$

The calculation of the SOAP power-spectrum vectors was carried out as implemented in the QUIP/quippy and GAP code (https://github.com/libAtoms/QUIP). The hyperparameters (cutoff r_c and broadness σ_{at}) of the SOAP kernel need to be carefully chosen to enable a proper distinction between crystal-like and amorphous-like atoms. As seen in Fig. 3.20, the distribution of $\overline{k}$ similarity of crystal-like and amorphous-like atoms was calculated with respect to RNG, and an overlapping distribution was found for moderate to large r_c values (< 8.5 Å). This suggests that a longer-ranged SOAP kernel is needed to distinguish crystal-like and amorphous-like regions in GST. The results were found to converge when using r_c larger than 9.5 Å; hence, $r_c = 9.0$ Å is used. Likewise, too small (< 0.3 Å) or too large (> 0.4 Å) settings for σ_{at} lead to a mix-up of crystal-like and amorphous-like particles (Fig. 3.20), and thus σ_{at} was set $\sigma_{\mathrm{at}} = 0.3$ Å.

The crystallization process of model 1 is shown in Fig. 3.21a with the SOAP crystallinity $\overline{k}$ indicated by colour coding. Both amorphous-like particles and local structural distortions can well be distinguished from the crystal-like atoms based on the SOAP crystallinity (Fig. 3.21a). More importantly, some atoms (colour coded in *black*), though seemingly crystal-like in the crystalline matrix, attain relatively low $\overline{k}$ values – indeed, these are antisite defects in the recrystallized model (Fig. 3.21b). This observation emphasizes that the SOAP crystallinity measure $\overline{k}$ can reveal chemical order in addition to the geometric order characterized in Fig. 3.16. The distribution of $\overline{k}$ for crystal-like (c-), amorphous-like (a-), and antisite atoms (Fig. 3.20) evidences that these three types of atoms can be well distinguished based on $\overline{k}$. The three peaks (labelled as 1, 2, and 3) in the $\overline{k}$ profiles also match the characteristics of c-, a-, and antisite atoms, respectively (Fig. 3.21c). These observations suggest that $\overline{k}$ can serve as a good indicator of both geometric and chemical order, used here in characterizing the crystallization process of GST.

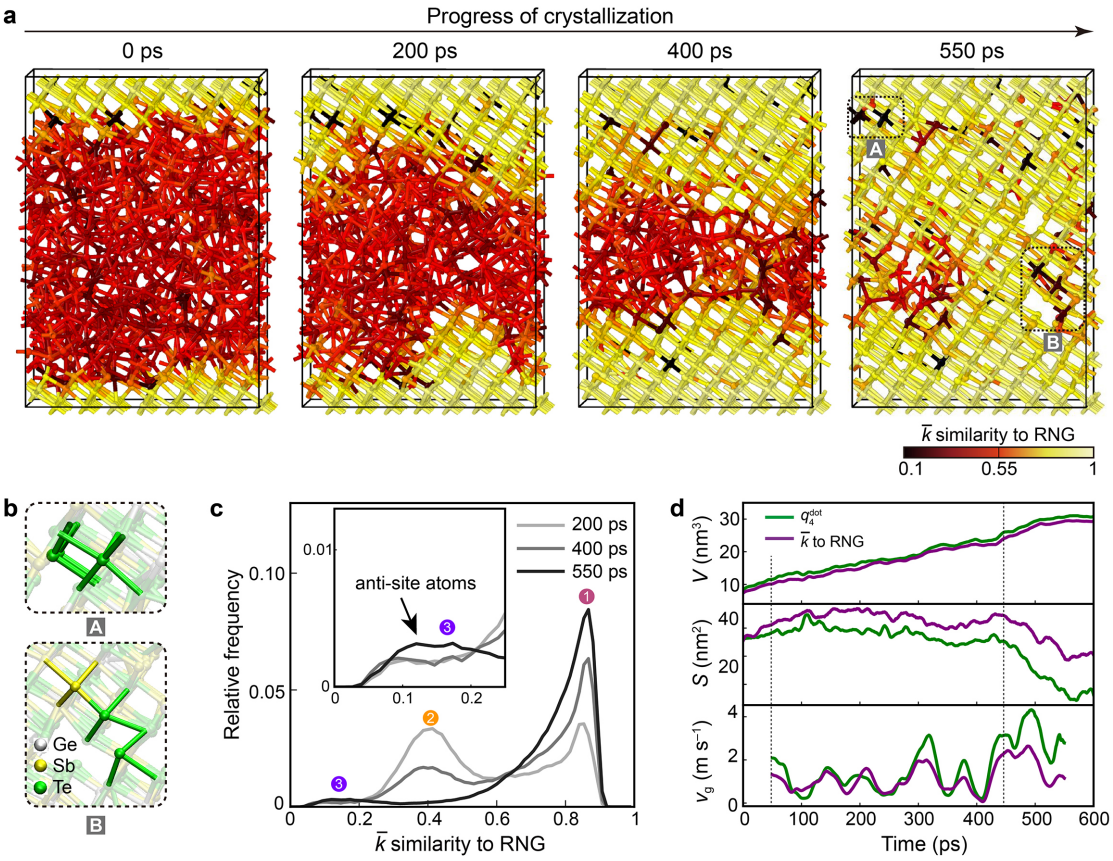


Figure 3.21: SOAP-based characterization of the crystallization process. (a) Snapshots of the crystallization simulation for model 1, with the $\overline{k}$ similarity (with respect to RNG) indicated by colour coding. The atoms in black are identified as antisite defects in the recrystallized model. (b) Enlarged atomic images for the areas labelled as "A" and "B" on the right-hand side of panel (a), containing antisite defects. (c) The $\overline{k}$ profiles of atoms in the 200-ps, 400-ps, and 600-ps configurations, in which the three peaks (labelled as 1, 2, and 3) match well with the distributions of a-, c- and antisite atoms in panel (b). (d) The volumes of the crystalline region V, areas S of the interface between the amorphous and the crystalline region, and crystal growth velocities v_g as a function of time for model 1, calculated using SOAP crystallinity $\overline{k}$ (*purple* lines) in contrast to q_{dot}^4 (*green* lines). A correction for "noise" atoms is performed before calculating these quantities. Volume and area data are smoothed using a Savitzky–Golay filter with a time window of 20 ps. The time window between two dashed lines in (d) is used to calculate the growth velocity.

As shown in Fig. 3.21d, with corrected "noise" atoms, the volume V of crystalline atoms, the area S of the crystal–amorphous interfaces, and thus the growth velocity v_g for model 1 were calculated. Based on the analysis on the SOAP crystallinity $\overline{k}$ (with respect to RNG), the results were compared to the q_{dot}^4 case (Fig. 3.21d). Note that the crystalline volumes calculated based on SOAP crystallinity are almost the same as those obtained using q_{dot}^4. However, the computed areas

of crystal–amorphous interfaces are larger when using SOAP, suggesting slightly rougher growth interfaces if the analysis is based on SOAP crystallinity. The estimated growth velocity, averaged over two models, is ≈1.2 m s^{-1} as calculated using SOAP, slightly smaller than that calculated by q^4_{dot}, 1.5 m s^{-1} (Fig. 3.18). Overall, the results are in good agreement with previous computational work on $Ge_2Sb_2Te_5$ (1 m s^{-1} at ≈600 K) [109] as well as the estimated value obtained from ultrafast differential scanning calorimetry experiments [281].

In short, the SOAP-based crystallinity analysis reveals the evolution of the antisite defects during crystallization in a quantitative way, enriching the observation of structural changes with chemical information. Hence, the SOAP-based crystallinity is used in the studies of crystallisation of various material systems in the following chapters.

Chapter 4

Tellurium-based phase-change selector materials

This chapter focuses on the elemental tellurium, a phase-change selector material, and demonstrates how ML potentials are used in solving fundamental textbook questions and practical problems in selector devices. Part of this chapter is based on a previously preprinted work on arXiv (Ref. 282). Some texts and figures have been reused. I would like to acknowledge the great contribution of Daniel F. Thomas du Toit for carrying out the hyperparameter optimisation and fitting the ACE potential. I would also like to thank Stephen R. Elliott, Wei Zhang, and Volker L. Deringer for their guidance and useful discussions for this work.

4.1 Chirality in elemental tellurium

Chirality, or handedness, is found in many organic molecules and also in inorganic solids: elemental selenium and tellurium, the common quartz form of silica (α-SiO_2), and cinnabar (α-HgS), which all crystallise in the same chiral space group. The ambient form of tellurium, α-Te, is among the simplest inorganic systems for studying chirality [283]. Two types of chirality are observed in α-Te, at different length scales: morphographic and crystallographic chirality. Morphographic chirality refers to the particle shape and is present in a wide range of organic and biological systems [284–286] as well as in inorganic materials [287–290]—indeed, its formation in Te nanoparticles was recently suggested to be driven by screw dislocations during crystallisation, as supported by atomic-resolution images from scanning transmission electron microscopy (STEM) [291]. Crystallographic chirality arises from the helical arrangement of atoms in the $P3_121$ (right-handed) or $P3_221$ (left-handed) structure, which is of interest because of its direct connection with physical proper-

ties. An example is the interplay of crystallographic chirality and electron spins in α-Te, which allows for the generation of spin polarization [292] and has been used in emerging spintronic devices [293, 294].

In addition to chirality-related applications, low-dimensional Te nanostructures (e.g., nano-tubes, -flakes, and -wires) have been used in functional devices [295, 296]. To obtain the desired morphography, chemical synthesis of these nanostructures can be controlled via nucleation and growth under different conditions [297–299]. Te has a very strong crystallisation tendency: the liquid element will quickly, even explosively, crystallise into helical structures at ambient temperature [299] or below [300]. A recent study reported a crystallisation time of less than 20 nanoseconds in as-deposited amorphous Te thin films [301].

Given a large electrical contrast between the crystalline (high-resistance) and liquid (low-resistance) phases [302], Joule heating pulses are used to melt crystalline Te films (the OFF state) and switch to the ON state. After programming, the OFF state is regained within nanoseconds, owing to the fast and spontaneous crystallisation. This volatile switching behaviour was shown to result in a non-linear current response in Te, making it a promising candidate for selector devices [301]. Selector devices are a critical component in three-dimensionally stacked crossbar arrays, which can suppress the leakage current and enable high-density memory applications [42–44]. Despite the simple monatomic composition which avoids compositional segregation, Te-based selectors still showed considerable fluctuations (noise) in switch-on voltages. A more in-depth understanding of the crystallisation processes in those devices is therefore required.

However, no real-space, atomic-level images have been reported for the crystallisation of Te that is at the scale of tens of nanoseconds [301], likely due to the simultaneous spatial and temporal resolution limits in experiments. Moreover, the atomic-level determination of handedness using microscopy makes it necessary to rotate the zone axes of the two-dimensional images around the spiral axis of Te chains

[283, 303]. Such complex experimental operations would be exceedingly difficult to realize.

To reveal the atomic-level mechanisms of crystallisation and chirality formation in α-Te, quantum-accurate atomistic simulations over millions of MD steps were performed, directly comparable to the length (10–20 nm) and time scales (10–20 ns) of switching operations in Te-based selectors [301]. To do so, a bespoke machine learning model was developed which reaches the accuracy of expensive meta-generalized gradient approximation (meta-GGA) density-functional-theory (DFT) computations, but at much lower computational cost. Based on ML-driven MD simulations, the short-lived cubic-like motifs were observed, which are key transient configurations in the crystallisation of Te and in the formation of left- or right-handed chiral chains. Further annealing at higher temperatures showed grain-boundary ordering and chirality transfer, also mediated by disordered cubic-like motifs. Based on these atomic-scale simulations, the switching mechanism of Te-based selectors can be explained.

4.2 The role of transient cubic-like motifs

The origin of elemental structures, including that of Te, has been discussed in the context of Peierls distortions: starting from a simple cubic structure, distortions and bond-breaking formally lead to different structure types [304, 305]. Such three-dimensional Peierls distortions give n short bonds and m long bonds (with $n + m = 6$), yielding the ground-state structure types of many pnictogen and chalcogen elements (Fig. 4.1). Unlike layered rhombohedral antimony (Sb) and black phosphorus (P), trigonal Te adopts a highly anisotropic and chiral crystal structure, which can be regarded as a distorted cubic phase with a 2 + 4 bonding environment (Fig. 4.1): each atom takes part in two strong covalent bonds, forming the helical chains [304].

An energy indicator is defined, i.e., the "cubic-phase formation penalty", E_{p}, expressed as the energy cost from a ground-state structure to a (hypothetical) simple-

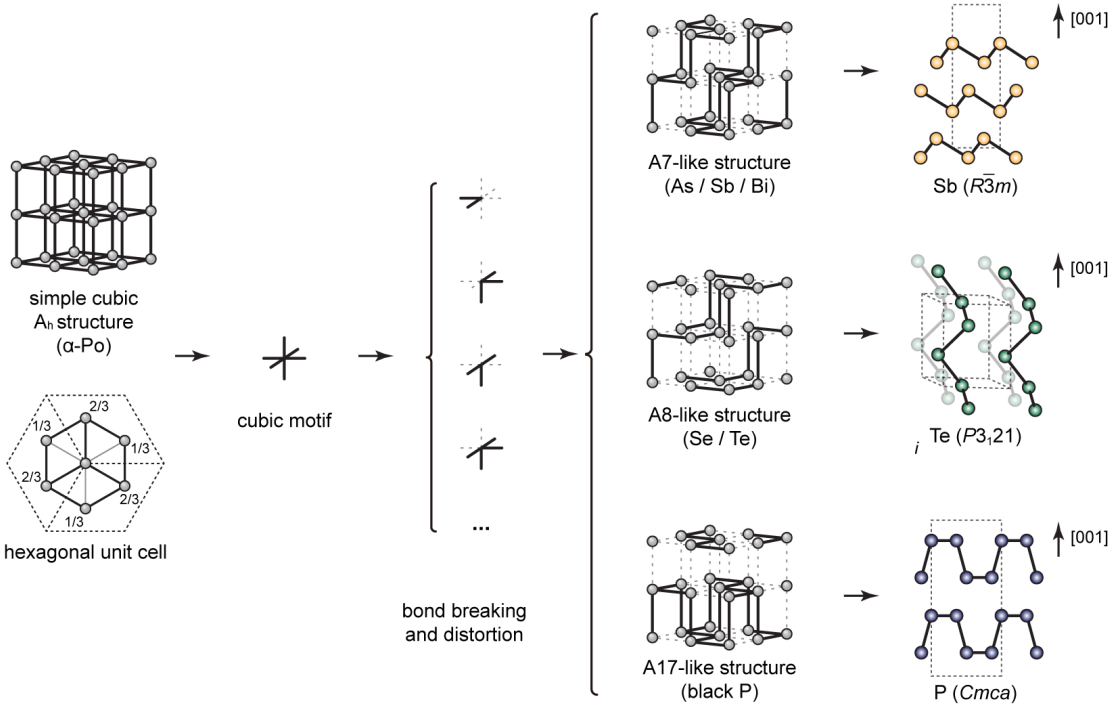


Figure 4.1: Three-dimensional Peierls distortion and the role of transient cubic-like motifs. Transitions from a simple cubic structure to ground-state Sb ($R\bar{3}m$), Te ($P3_121$), and P ($Cmca$) structures via bond breaking and distortion.

cubic structure (i.e., E_p =$E_\mathrm{cubic\text{-}phase}$ -$E_\mathrm{ground\text{-}state}$), for various pnictogen/chalcogen elements and chalcogenides (Fig. 4.2). Such a quantified indicator inherently measures the energetic difficulty of forming a cubic-like phase. A gap at $E_\mathrm{p} \approx 100$ meV atom^{-1} clearly separates the lighter, third- and fourth-period elements (i.e., P, As, Se) and the heavier, fifth- and sixth-period elements (i.e., Sb, Bi, Te, Po). Notably, these two groups of elements have a different crystallisation tendency. For the period-3 / -4 elements with $E_\mathrm{p} > 100$ meV atom^{-1}, their crystallisation temperatures are highly above room temperature [306–308], and the simple cubic phase was only observed under high pressures, e.g., in P [309], As [310], and Se [311]. By contrast, the period-5 / -6 elements with $E_\mathrm{p} < 100$ meV atom^{-1}, e.g., Sb [145], Bi [312], and Te [299], show ultrafast crystallisation, and can only remain in an amorphous state under nanoscale confinement [145, 299, 312]. It was also suggested that Sb formed a disordered cubic phase during fast nucleation in AIMD simulations [313, 314]. Moreover, many GST phase-change memory materials, with E_p values of less than

100 meV atom^{-1} [11, 315], first crystallise into a metastable cubic phase within nanoseconds and then further into a stable hexagonal phase on high-temperature annealing [316]. Hence, it is inferred that the easier formation of a cubic-like phase leads to faster crystallisation.

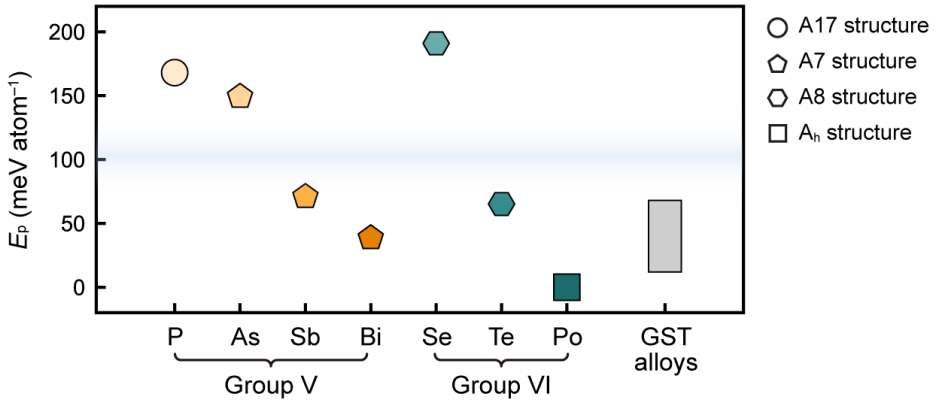


Figure 4.2: The calculated cubic-phase formation penalty, E_p, defined as the energy difference between the hypothetical cubic phase and the ground-state phase for various group-V / -VI elements and compounds. A gap at $E_p \approx 100$ meV atom^{-1} clearly separates the lighter, third- and fourth-period elements (i.e., P, As, Se) and the heavier, fifth- and sixth-period elements (i.e., Sb, Bi, Te, Po).

To evaluate the energy cost of forming a hypothetical cubic Te phase, a grid search for the energetic landscape of different atomic configurations was performed (Fig. 4.3): the symmetry and lattice parameters of the DFT-relaxed hexagonal unit-cell of trigonal Te were fixed, and the relative positions of three atoms in the unit-cell were tuned. The three fractional coordinates are (0, 0, 0), (x, y, 1/3), and (y, x, 1/3), with x, y $\in$ (0, 1) and x $\geq$ y. Based on the DFT-calculated energy profile, an energy barrier of less than 90 eV atom^{-1} was found for forming a cubic-like motif. In addition, direct transfer routes between left- and right-handed Te—a "tunnel" was found that bypassed the idealized cubic phase and connected two chiral configurations (Fig. 4.3). Along the route, no saddle points were observed, suggesting that disordered cubic-like Te is not a metastable phase but merely a transient form.

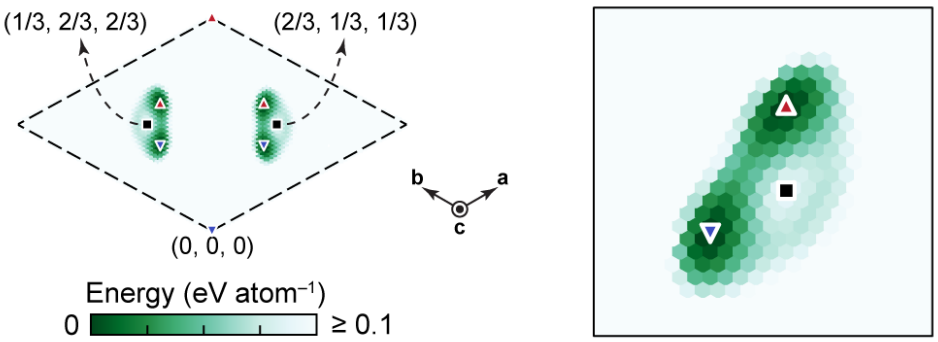


Figure 4.3: The energy profile of the projection along the c-axis, showing the possible transition between the hypothetical cubic phase (black square) and two energetically equivalent trigonal phases, i.e., left-handed (red triangle) and right-handed (blue inverted triangle) Te. The result was computed using DFT. The zoom-in figure is shown in the right panel.

A possible pathway to the chirality transfer in elemental tellurium is shown in Fig. 4.4. It is notable that simply rotating or displacing the right-handed ($P3_121$) α-Te crystal without breaking chemical bonds cannot give the left-handed ($P3_221$) form, and vice versa. However, such chirality transfer is theoretically feasible by using the cubic phase as an intermediate, rearranging into a new $2 + 4$ bonded motif that gives rise to a different chiral helix (Fig. 4.4). Such a transfer only requires very few atomic displacements, and the computed low energy barrier supports the idea that chirality transfer via cubic-like motifs might happen at finite temperatures (Fig. 4.3). No saddle points were found between the ground state and a hypothetical cubic Te phase, suggesting that disordered cubic-like Te is energetically unstable (i.e., not a metastable phase).

4.3 Fitting ML potentials for tellurium

To study the crystallisation of supercooled amorphous Te, ML potentials were utilised to run large-scale MLMD simulations over tens of nanoseconds. The development and validation of the ML potential model were based on a domain-knowledge-based iterative training protocol [180] and the atomic cluster expansion framework [172] with optimized hyperparameters [207].

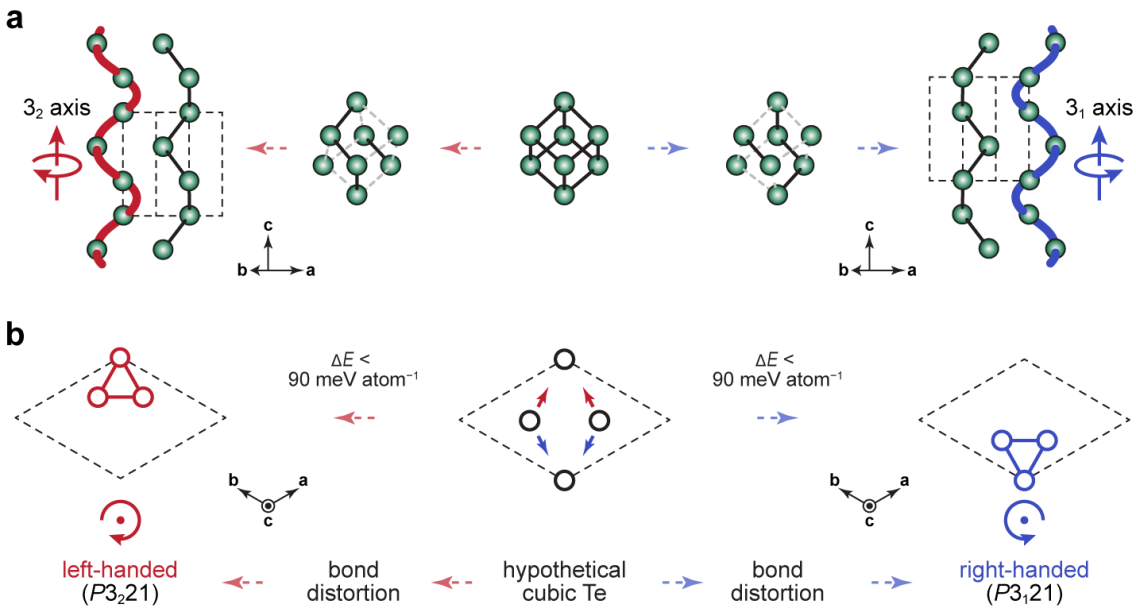


Figure 4.4: The microscopic pathway to chirality in Te crystals. (a) The transitions from a simple cubic structure to the left-handed ($P3_221$) and right-handed ($P3_121$) crystal structure of Te, via bond breaking and distortion, sketched along the [120] direction. Dashed lines represent the unit cell of α-Te. The chiral helices formed by the left-handed (3_2) and right-handed (3_1) screw axes are emphasised by thick red and blue lines, respectively, drawn in a style similar to Ref. 283. (b) Top view of the crystallographic unit cells along the c-axis. The computed energy barrier of such transitions is less than 90 meV atom^{-1} (cf. Fig. 4.3).

4.3.1 Choice of fitting frameworks

Firstly, the Gaussian approximation potential (GAP) framework was used to fit an ML interatomic potential for tellurium. The training dataset was generated using a domain-knowledge-based iterative protocol. It has been shown in the previous work [180] and will be discussed in the next chapter that such a protocol produced reliable ML potentials, enabling the accurate description of chemically and structurally complex GST alloys under realistic device conditions. Highly converged DFT calculations were performed to obtain the energies, forces, and stresses of the training structures (i.e., for the labelling of the reference dataset). The meta-generalised gradient approximation (meta-GGA) type Tao–Perdew–Staroverov–Scuseria (TPSS) functional [317] was employed, which has been demonstrated to enable accurate AIMD simulations of liquid and amorphous Te [318, 319]. To further accelerate the MD simulations, the ACE framework [172] was employed, which is faster than GAP

by at least two orders of magnitude [182]. An ACE model was then fitted based on the reference dataset of the GAP model using XPOT's interface [207, 208] to PACEmaker [172, 174, 182]. Only per-structure energy and per-atom forces were used in the fitting of the ACE potential. For this work, the hyperparameters were optimised for a related material from an ongoing project that has not yet been published (Zhou et al., in preparation). This results in a non-linear model with 3000 functions and three embeddings of the form:

$$E_i = \varphi + 4.6\sqrt{\varphi} + 0.1\varphi^2$$

where E_i is the energy of atom i and φ is the atomic property.

The ACE model was fitted without any structural weighting, i.e., each structure has the same weighting in the loss function minimization during fitting. An upfitting approach for fitting the ACE potential was used. This consists of two separate fits of the potential, where the initial fit focused on converging the forces using κ=0.8, and the second fit (upfit) improved the energy predictions by using κ=0.02, where κ controls the relative weighting of the energy and force errors in the loss function. The protocol used here is an updated version of the upfitting protocol used in Ref. 207. The initial fit used a power order ladder fitting protocol, as defined in Ref. 174, with a ladder step of 1000 functions, for a total of 2250 (750 $\times$ 3) iterations. The upfit did not use ladder fitting (this is not possible when fitting atop an existing potential) and lasts 2000 iterations.

Both models provided accurate results as compared to DFT references and experimental data. Notably, those ML models can capture the energetic difference between cubic-like and chain-like motifs (Fig. 4.5), with the largest predicted error being less than 20 meV atom^{-1} as compared to DFT. The accurate reproduction of the transition profile between cubic-like and chain-like motifs using ML potentials is key to revealing the important role of transient cubic-like motifs in MLMD simulations, otherwise the subtle structural changes can barely be captured.

All atomistic simulations here were performed using the ACE potential, due to an improved computational efficiency of ACE. In contrast, the GAP model was used to compute the total energy and per-atom energies of the structural models, with the help of the Python package QUIP and its Python interface quippy, because GAP models have been successfully shown to provide insightful energetic information in complex elemental systems, e.g., in silicon [279].

4.3.2 Domain-knowledge-based iterative training

To generate a reliable reference dataset covering a wide range of structural diversity in Te, a domain-knowledge-based iterative training protocol was used. To construct an initial reference dataset, which was called "iter-0", the isolated Te atom and Te–Te dimer configurations (in a large box of $20 \times 20 \times 20$ nm^3) were first added in the iter-0 database. Next, in total 8 rounds of GAP-based random structure search (GAP-RSS) processes were carried out to sample a wide range of random structural space [197, 202]. The unit cells and supercells of crystal structures from the Materials Project database were also included [320], as well as various copies of these crystal structures with disordered lattice parameters and atomic coordinates. In addition, the configurations of liquid, supercooled liquid, and amorphous Te, as well as intermediate configurations during the growth simulations were included in the iter-0 database, which were all obtained from AIMD simulations. Starting from the iter-0 dataset, a series of iterative training was carried out for GAP potentials, which can be divided into: (i) standard iterations that contain abundant supercooled liquid Te configurations at different temperatures and with varying densities; and (ii) domain-specific iterations which include intermediate transition states during melting and crystallisation of Te. The domain-specific iterations were specifically designed in this protocol so that the ML model can well describe the phase-transition processes of Te. In total, 30 standard iterations and 20 domain-specific iterations were carried out in the iterative training process.

To obtain a stable GAP potential, a screening of the original training dataset

was carried out and "unreasonable" configurations were ruled out, including: (i) bulk structures with too small or too large densities (i.e., $\rho < 0.015$ or $\rho > 0.045$ Å^{-3}); (ii) structures with too high cohesive energies ($E_{\mathrm{coh}} > 20$ eV per structure); and structures with too large force components ($|F_{ij}| > 20$ eV Å^{-1}, where i runs over all atoms in the structure and j represents force components along the x, y, and z directions). Note that most of the removed configurations are high energy RSS structures, in which very short atomic distances were found. By contrast, to fit a stable ACE potential using the same reference database, the aforementioned "unreasonable" configurations that were removed from the GAP model were kept, but all Te–Te dimer configurations were removed. Note that retaining these "unreasonable" configurations improved the robustness of the ACE potential without significantly reducing accuracy on lower energy structures, and removing the dimers (i.e., rather open structures with very low atomic densities) increased prediction accuracy on the bulk Te testing dataset. More importantly, unlike GAP, ACE potentials do not have an explicit, separate two-body term, reducing the effectiveness of adding dimer configurations. A detailed summary of the training datasets for GAP and ACE models is provided in Table 4.1.

Table 4.1: Dataset sizes for GAP and ACE models. The table lists the number of cells and atoms for various types of data used in the models.

Type	GAP		ACE	
	Cells	Atoms	Cells	Atoms
Free atom	1	1	0	0
Dimer data	33	66	0	0
RSS structures	684	7718	800	9054
Crystalline phases	284	6081	304	6673
AIMD configurations	117	28788	117	28788
Standard iterations	395	71100	395	71100
Domain-specific iterations	317	69840	317	69840
Total	1831	183594	1933	185455

To label the reference dataset, the energies and forces were computed from single-point calculations using VASP [262, 263]. Projector augmented wave (PAW) pseu-

dopotentials were used [261, 262]. The meta-GGA type TPSS functional [317] was employed, which has been demonstrated to enable accurate AIMD simulations of liquid and amorphous Te [318, 319]. The cut-off of the plane-wave expansion was set to 650 eV. In the single-point calculations, we used an energy convergence criterion of 10^{-7} eV per atom. A k-point grid with a spacing of 0.2 $Å^{-1}$ was used to sample reciprocal space.

4.3.3 Validation of ML potentials

All of the ML potentials were carefully validated, following guidelines of validating ML potentials in Ref. 209. A numerical validation was first carried out by computing the prediction errors based on an external dataset that was not included in the training reference dataset. This external testing dataset contains different kinds of structures of Te, including melt-quenched disordered forms (viz. the liquid, supercooled liquid, and amorphous states) at different temperatures, and the intermediate configurations taken from the AIMD simulations of the templated growth processes. Large numeric errors of the external testing dataset might suggest an overfitting to the training database and thus an inaccurate ML potential model. In the numeric tests, the small root-mean-square errors (RMSE) of both predicted energies and forces were obtained, which indicated a good numerical accuracy of both GAP and ACE models (Table 4.2).

Table 4.2: The computed root-mean-square error (RMSE) values of a test dataset that was not included in the training data of the ML models. The structures in the test dataset can be split into four different categories, including liquid, supercooled liquid, amorphous structures of Te and intermediate configurations of the growth process, which were taken from AIMD simulations.

	GAP		ACE	
	RMSE energies (meV $atom^{-1}$)	RMSE forces (meV $Å^{-1}$)	RMSE energies (meV $atom^{-1}$)	RMSE forces (meV $Å^{-1}$)
Liquid	13.8	164	8.1	124
Supercooled liquid	41.9	169	30.4	132
Amorphous	52.2	164	38.3	131
Growth frames	26.6	142	16.7	107

Next, more comprehensive validations were carried out based on physical properties of Te. Firstly, the energy profiles of trigonal Te were computed using both GAP and ACE models, which were then compared to the DFT reference, as shown in Fig. 4.5. Notably, the energy difference of subtle structural changes between cubic-like and chain-like configurations is key to describing crystallisation and chiral transfer in Te, which should be accurately captured by the ML potentials. Based on the plot energy profiles using DFT, GAP, and ACE models, it is clear that the energy difference between cubic-like and chain-like motifs can be well captured by both the GAP and ACE models, giving similar results as compared to the reference DFT. Only a small numeric discrepancy (less than 20 meV atom^{-1}) was observed. However, the transition route between left- and right-handed Te (indicated by the "tunnel" between the points of two chiral fragments) was well captured by the ML models, which is consistent with the DFT results.

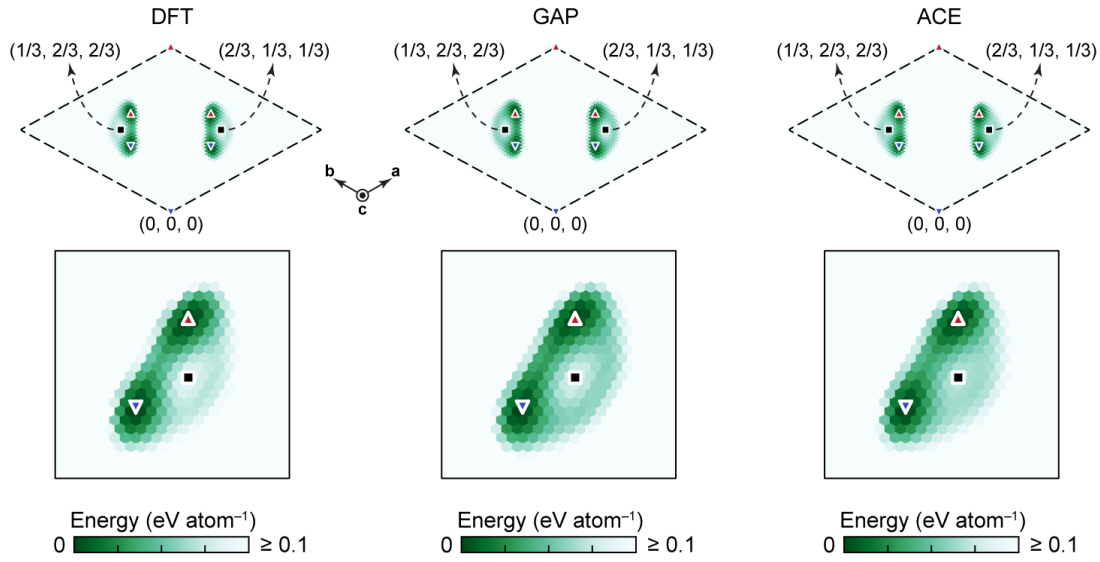


Figure 4.5: The energy profile of the projection along the c-axis, showing the possible transition between the hypothetical cubic phase (black square) and two energetically equivalent trigonal phases, i.e., left-handed (red triangle) and right-handed (blue inverted triangle) Te. The results were computed using DFT, GAP potential, and ACE potential were compared, showing good agreement with each other.

In addition, more structural validations were carried out on liquid, supercooled liquid, and amorphous Te structures obtained from both AIMD and MLMD simula-

tions. The radial distribution function (RDF) and the angular distribution function (ADF) were first calculated (Fig. 4.6a and b). The RDF is a general indicator for the medium-range order of liquids and amorphous Te, while the ADF is a more sensitive indicator for the local fragments in disordered Te. The overall good agreement between the reference AIMD and the MLMD simulations was observed, and only a very small discrepancy was found at the first peak of the RDF of the amorphous Te at 300 K (Fig. 4.6a).

Moreover, the angular-limited three-body correlation (ALTBC) function was computed, which is a quantified indicator for the Peierls distortion in liquid, supercooled liquid, and amorphous Te. The ALTBC function expresses the probability of having a bond of length r_1 almost aligned with a bond of length r_2 (with angular deviations smaller than 30°). It is found that the results from both GAP- and ACE-driven MD simulations were in good agreement with the AIMD reference, indicating that both ML models can accurately describe the formation of Peierls distortion in liquid, supercooled liquid, and amorphous Te (Fig. 4.6c).

Lastly, the computed structure factors of liquid Te obtained from MLMD simulations at various temperatures were compared to experimental references. Note that small discrepancies were already observed in two different experimental results, and the computed structure factor from the MLMD simulations was overall in accordance with the experimental data of liquid and supercooled Te, indicating that the ML models can accurately describe the structures of liquid and supercooled liquid Te (Fig. 4.7). However, no experimental structure factor data was found for the amorphous Te, presumably due to the fast and simultaneous crystallisation of Te at the room temperatures, making it hard to maintain the amorphous form for subsequent measurements.

4.4 Crystallisation of Te at the atomic scale

Large-scale (102,000 atoms in a 15.2 × 15.5 × 15.0 nm^3 box) MLMD simulations over tens of nanoseconds were carried out to study the crystallisation of supercooled

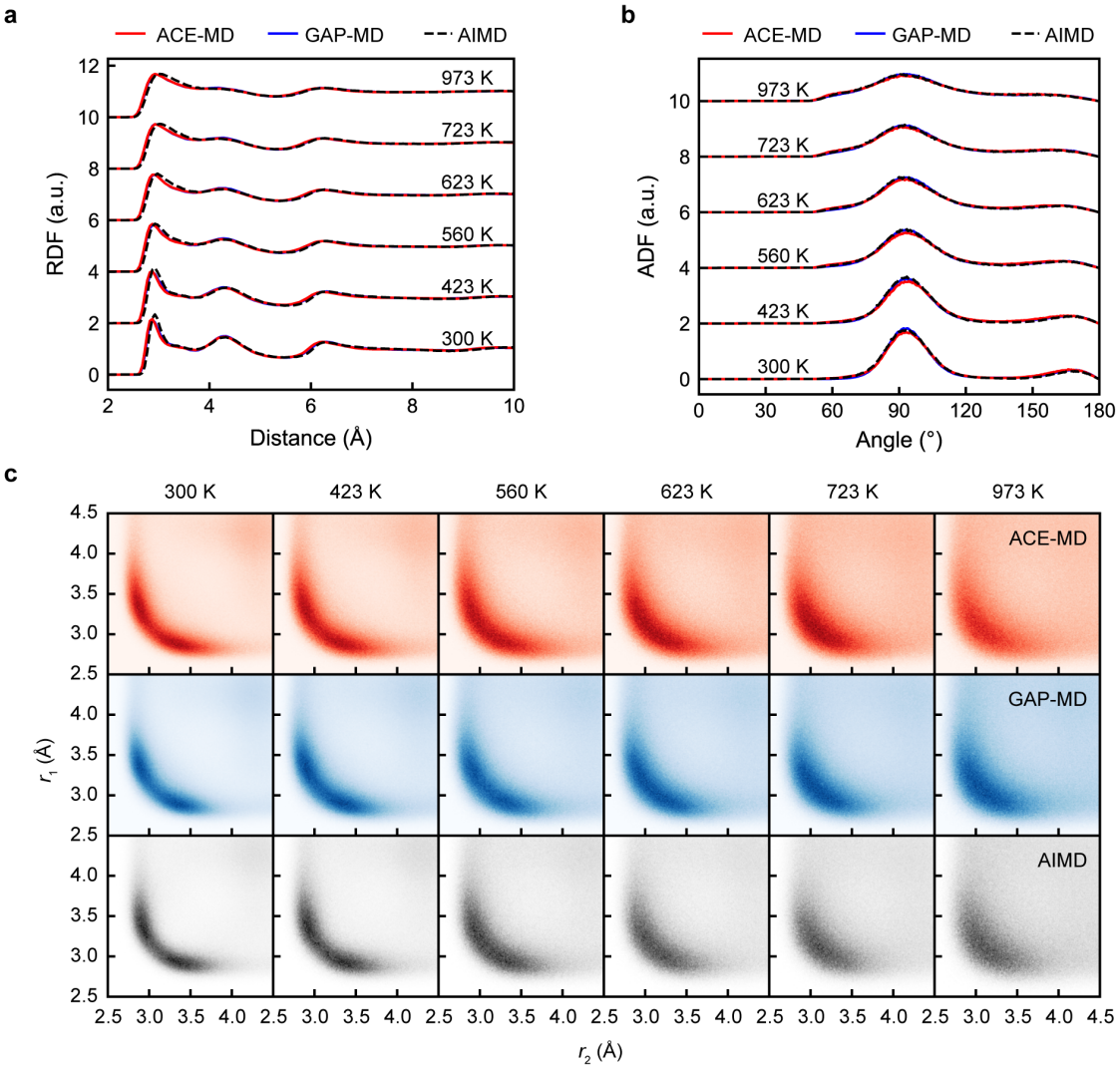


Figure 4.6: Structural validations for ACE (red) and GAP (blue) potential models as compared to the AIMD reference (black). (a) Radial distribution function (RDF) computed at various temperatures. (b) Angular distribution function (ADF) at different temperatures. (c) Angular-limited three-body correlation (ALTBC) functions as indicators for the Peierls distortion in liquid, supercooled liquid and amorphous Te at different temperatures. The ALTBC function expresses the probability of having a bond of length r_1 almost aligned with a bond of length r_2 (with angular deviations smaller than 30°).

amorphous Te at 323 K, slightly above the room temperature. In all, five independent runs were performed for better statistics, and all five structural models crystallised into polycrystals (Fig. 4.8).

Note that all MLMD simulations shown in this chapter were carried out using LAMMPS [322] with an interface to the pacemaker package, which is available freely for academic research at https://github.com/ICAMS/python-ace. The canonical

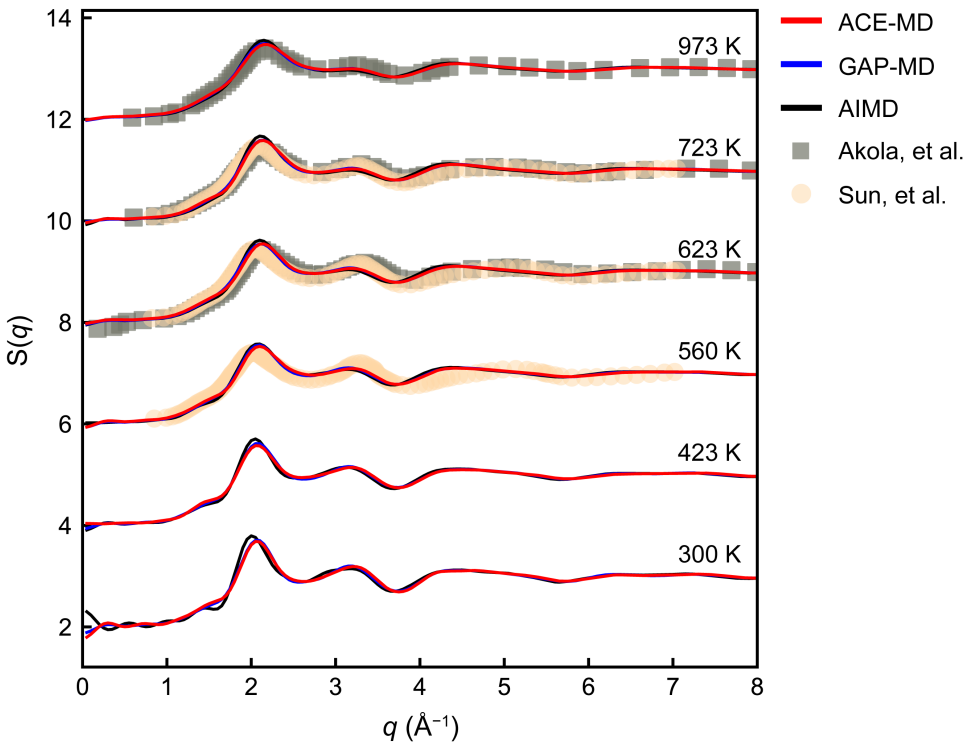


Figure 4.7: Comparison of structure factors of supercooled liquid and amorphous Te between simulated results (using ACE-MD, GAP-MD, and AIMD) and experimental data. Data for different temperatures were considered, including 300, 423, 560, 623, 723, and 973 K. Experimental data were taken from Akola, et al. (Ref. 318) and Sun, et al. (Ref. 321).

ensemble (i.e., constant volume and temperature; NVT) was used in all ML-driven MD simulations with a Langevin thermostat (using the default damping value in LAMMPS). In this regard, the density of the structural model was fixed, which represents the restrained geometrical conditions imposed by the surrounding materials in confined Te-based selector devices. The timestep was set to 1 fs for all ML-driven MD simulations.

4.4.1 Cubic- and chiral-like motifs

Figure 4.9 shows more details of one such crystallisation simulation. The smooth overlap of atomic positions (SOAP) similarity measure was used to quantify the degree of "per-atom crystallinity", $\overline{k}$ [114, 170], with respect to ideal α-Te. The details of this similarity measure have been discussed in Section 3.5 and in Ref. 114. Notably, such "crystallinity" indicator allowed to distinguish amorphous-like ($\overline{k}$ <

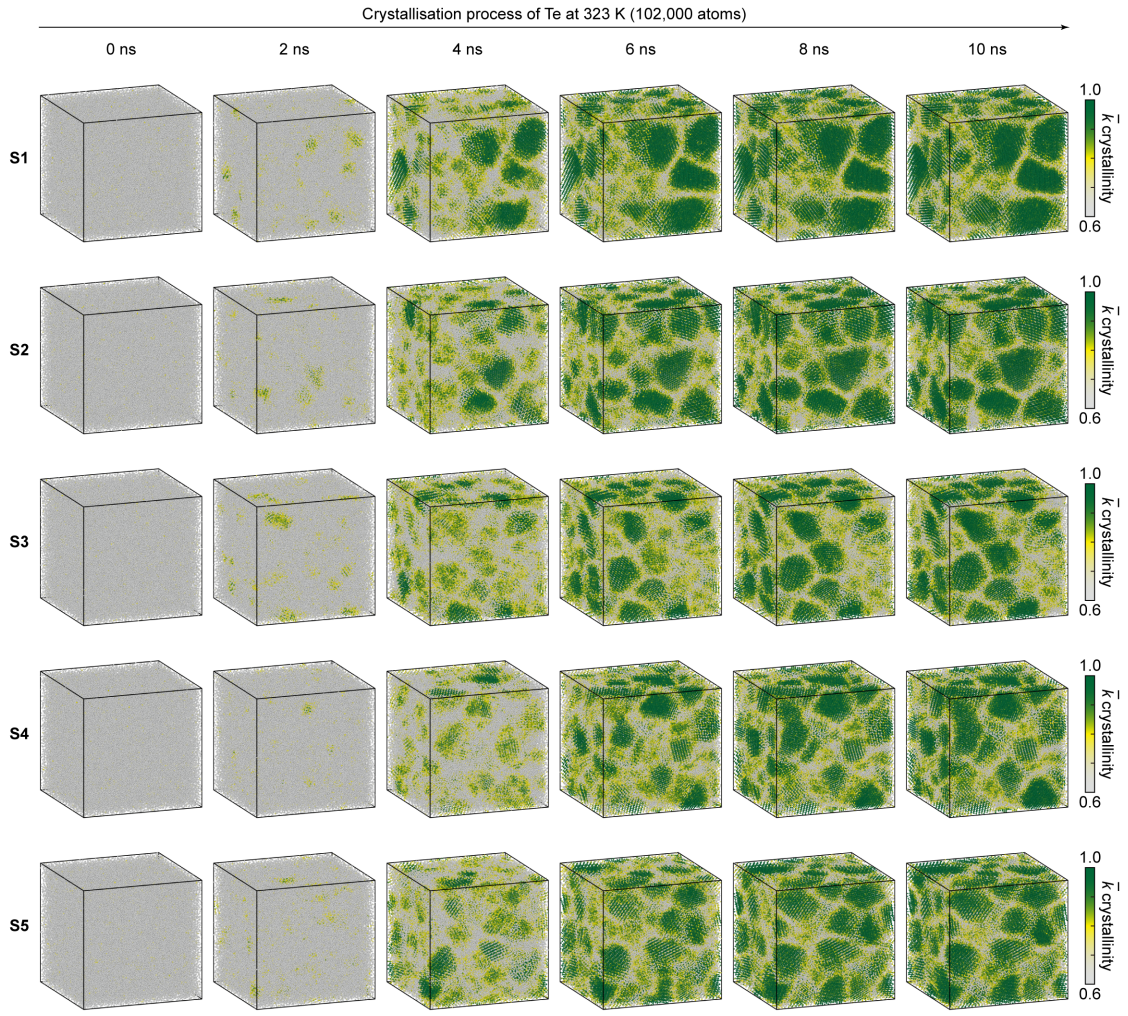


Figure 4.8: Five independent 10-ns crystallisation simulations at 323 K, labelled **S1** to **S5**. Atoms are colour coded using per-atom SOAP crystallinity $\overline{k}$, illustrating the gradual structural ordering from the (disordered) supercooled liquid phase to the crystallised polycrystal. The trajectory of model **S1** was shown in Fig. 4.9a for further illustration.

0.75) and crystal-like ($\overline{k} > 0.75$) atoms in the simulation. Moreover, the crystal-like atoms can be further separated into two types using the SOAP crystallinity: disordered cubic-like motifs ($0.75 < \overline{k} < 0.85$) that are three- or four-coordinated with bond angles close to 90°; and chain-like atoms ($\overline{k} > 0.85$) that clearly belong to either left- or right-handed helices. Both disordered cubic-like and chain-like atoms were found in supercooled amorphous Te [319, 323]. This indicates similar local bonding environments in the amorphous and crystalline forms.

4.4.2 Nucleation in Te

The start of the crystallisation process (≈1 ns into the simulation) was signalled by the creation of a crystal-like seed containing tens of the over-coordinated cubic-like atoms and no chain-like atoms (Fig. 4.9a). After that, more cubic-like seeds appeared, indicating a nucleation process with multiple nuclei, consistent with experimental observations at room temperature [299]. More comprehensive analyses were carried out: the potential energy dropped markedly (Fig. 4.9b) and the numbers of cubic-like atoms increased (Fig. 4.9c) during incubation (0–1 ns) and nucleation (1–3 ns). At ≈3 ns, the cubic-like precursors suddenly underwent swift and synergistic distortions, forming chain-like fragments with only two primary bonds (2 + 4) and near-103° bond angles (Fig. 4.9a). Such fast and collective behaviour was also evidenced by the abrupt increase of two-coordinated atoms at ≈3 ns (Fig. 4.9d). The estimated lifetime of the cubic-like precursors is ≈1 ns, based on the different onsets of forming disordered cubic-like and chain-like motifs (Fig. 4.9c inset).

The analyses of ADF and nearest neighbour further reveal the subtle structural change from cubic-like to chain-like fragments. As shown in Fig. 4.10, at the starting of nucleation, a large peak of ADF at 90° was observed, indicating the formation of cubic-like motifs. As the crystalline seeds grew, the peak at 90° became less pronounced, and the peak gradually moved towards 103°. A small peak at ≈180° also became less and less pronounced throughout the nucleation, indicating the disappearance of cubic-like fragments. In addition, the n-th nearest neighbour correlation functions also showed a clear transition from cubic-like motifs. The distributions of first and second nearest neighbours gradually separated from the rest of the neighbours, indicating the formation of 2 + 4 bonding environments.

4.4.3 Crystal growth in Te

Following nucleation, fast growth was observed from 3–5 ns, concomitant with a linear decrease in potential energy (Fig. 4.9b) and the increasing presence of crystal-

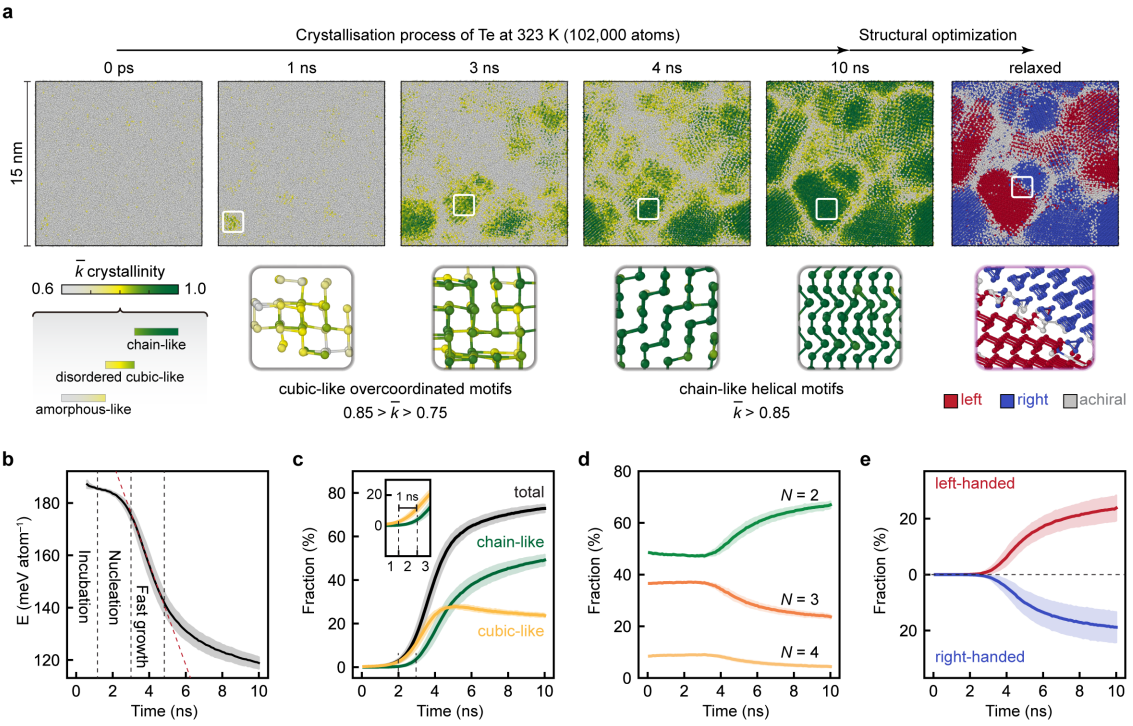


Figure 4.9: Nucleation and crystallisation of tellurium. (a) Crystallisation of supercooled Te from ML-driven molecular dynamics, at 323 K for 10 ns. The structural model contains 102,000 atoms. Atoms are colour-coded using the per-atom SOAP crystallinity $\overline{k}$, illustrating the gradual structural ordering from the (disordered) supercooled liquid phase to the polycrystal. Atoms in the rightmost panel are colour-coded according to chirality. The lower panels highlight typical local structural motifs during the crystallisation process. (b) Energy profile of the crystallisation process, divided into three different stages, viz. incubation, nucleation, and fast growth. (c) Evolution of the fraction of chain-like and cubic-like motifs. The inset shows different onsets for forming cubic-like and chain-like motifs. (d) The calculated fraction of atoms having various coordination numbers (N = 2, 3, 4, respectively) during the crystallisation process. (e) The counts of atoms in left- and right-handed helices formed during the crystallisation. All data in (b)–(e) are averaged over five independent samples (500 snapshots in total), and the shaded areas indicate the standard deviation for the 500 snapshots.

like environments (Fig. 4.9c). Notably, disordered cubic-like atoms still appeared at the crystalline–amorphous interfaces at the growth fronts, whereas atoms inside the grains assembled into chiral chains (Fig. 4.9a). To study more details about the growth process in Te, elongated crystal structural models (containing 720 atoms) were used, larger than recently reported AIMD simulations [324]. Three atomic layers of the crystalline atoms were pre-fixed in two crystal planes, i.e., (001) and (210) planes. A melt-quench process was used to eliminate the ordering in the remaining part of the models and thus generate the structural model for growth

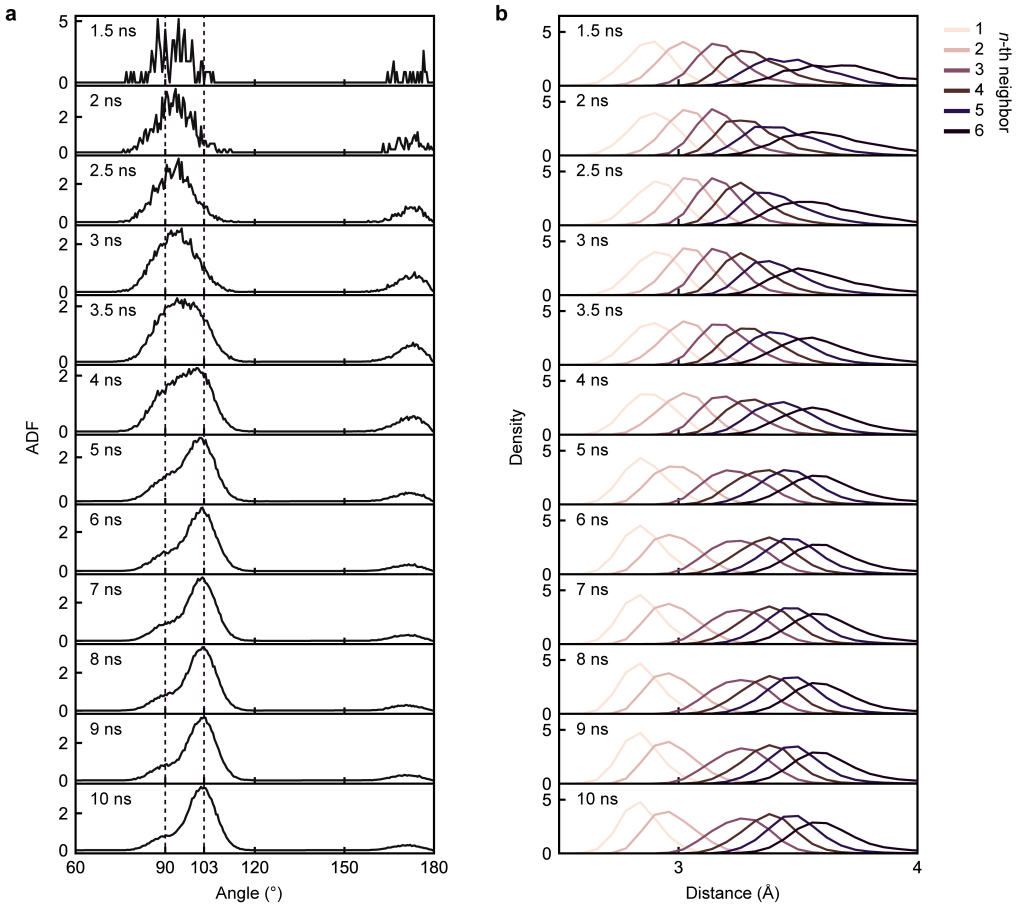


Figure 4.10: Local structural evolution of crystal-like atoms during the 10-ns crystallisation at 323 K. The calculated (a) angular distribution function (ADF) and (b) n-th nearest neighbour correlation functions show a clear transition from cubic-like motifs (with a coordination number exceeding two and with 90° bond angles, marked by the vertical dashed lines) to chain-like helical fragments (with 2 + 4 bonding environments and 103° bond angles).

simulations: the initial structural model was first randomized at 3,000 K for 30 ps; it was then cooled down to 1,000 K, above the melting point of Te, within 30 ps; the liquid model was equilibrated at 1,000 K for another 30 ps and then further quenched to 423 K. The fixed atoms were then released in the growth simulations at 423 K (Fig. 4.11).

All AIMD simulations shown in this chapter were performed using the "second-generation" Car–Parrinello scheme [325], as implemented in the Quickstep package of CP2K [326]. The Kohn–Sham orbitals were expanded using a triple-zeta plus polarization Gaussian-type basis set and we used plane waves with a cut-off of 300 Ry to expand the charge density. We used scalar-relativistic Goedecker pseudopotentials

[327] and the TPSS functional [317]. The Brillouin zone was sampled only at Γ point, and the timestep was 2 fs.

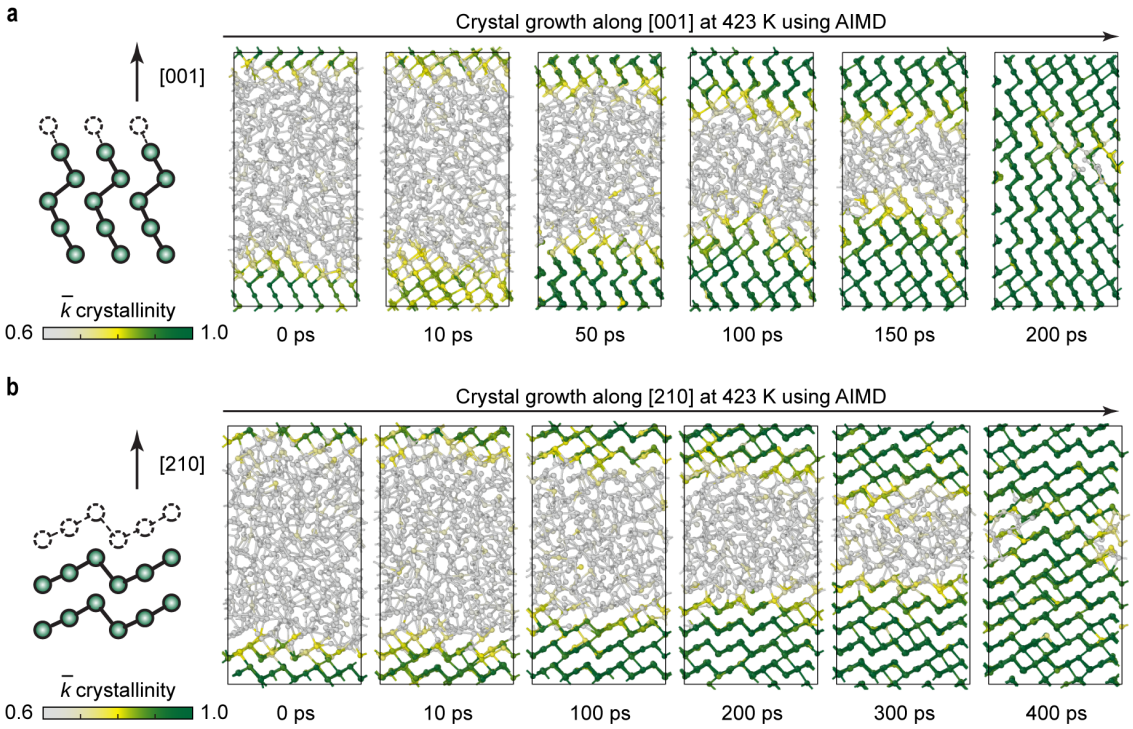


Figure 4.11: Growth simulations of a templated model (720 atoms) at 423 K using AIMD. (a) Growth simulation along the [001] direction, i.e., the chain direction of Te. (b) Growth simulation along the [210] direction, i.e., perpendicular to the chain direction of Te.

Notably, the cubic-like motifs facilitate inter-chain bonding, leading to nearly isotropic growth (rather than anisotropic growth along the chain direction only). However, the growth along the chain direction (i.e., [001] direction) was still faster than that perpendicular to the chain direction. In contrast to the sudden and collective transition from cubic- to chain-like fragments in nucleation, the growth proceeded atom-by-atom, similar to the growth mechanism of Ge–Sb–Te phase-change materials [109, 114]. The resultant polycrystalline model has ≈80 % of crystal-like atoms (Fig. 4.9c). Note that the recrystallised structure in the simulations is consistent with atomic-level images in experiments [301] and in recent AIMD simulations [324], but very different from the icosahedral and body-centred-cubic like clusters found in previous simulations based on an empirical force field [328].

4.4.4 Quantification of per-atom chirality

Based on the MLMD simulations, both left- and right-handed Te chiral fragments were found in the recrystallised structure. To quantify the per-atom chirality in the recrystallised Te structures, a structural indicator, the chiral index, was proposed to describe the well-defined chiral shape of any given atom. A similar idea of identifying chirality in water molecules was given in Ref. 329, which was based on the geometrical connections of neighbouring molecules. The determination of per-atom chirality can be separated into three steps (Fig. 4.12).

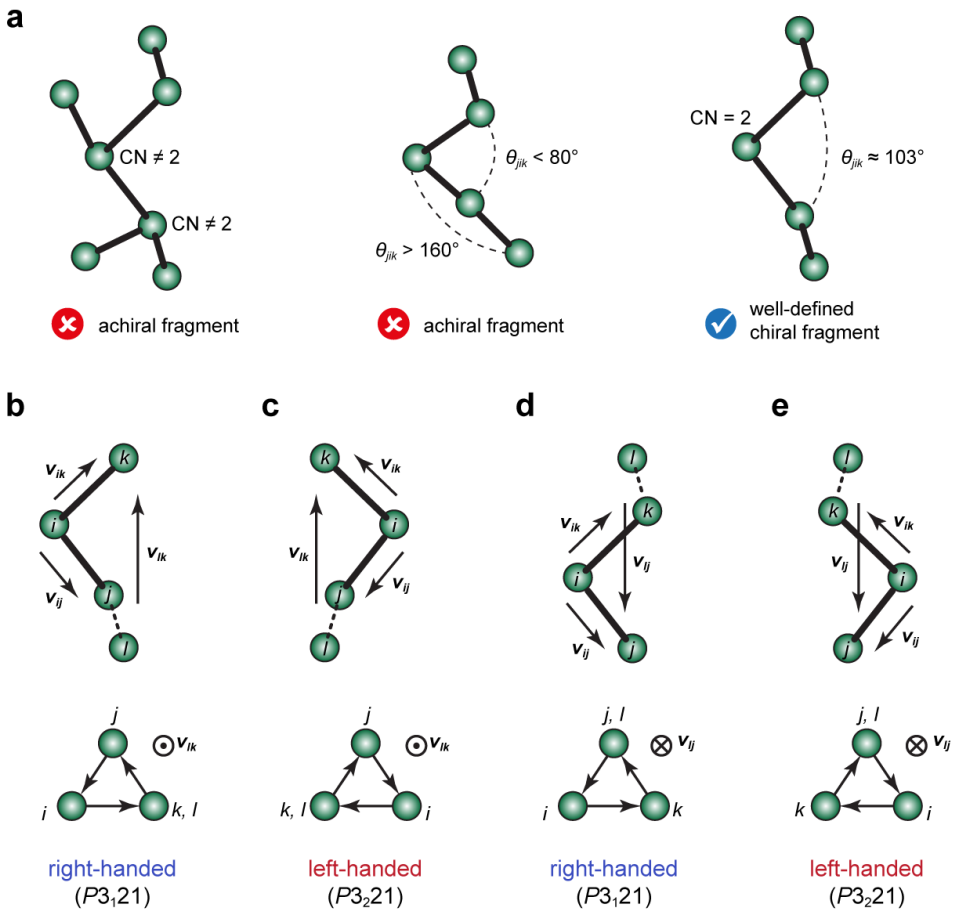


Figure 4.12: Atomic-level determination of per-atom chirality in Te structures. (a) The difference between achiral and well-defined chiral fragments, which can be distinguished based on the computed coordination number (CN) and bond angles (θ_{jik}). (b)–(e) Four different cases show the atomic-level determination of per-atom chirality based on the geometric connectivity of four bonded atoms in sequence.

Firstly, the per-atom crystallinity was computed for each atom. Note that only

the crystal-like atoms (i.e., with a high similarity to the reference α-Te) have a chance to be a "chiral atom", and therefore the disordered atoms are all achiral. Moreover, the coordination number (CN) of a well-defined chiral Te atom can only be two. Hence, crystal-like atoms with incorrect coordination numbers (i.e., CN $\neq$ 2) were regarded as achiral atoms. An atomic cutoff of 3.2 Å was used to determine the coordination numbers (Fig. 4.12a).

Secondly, the connection of a given atom i to its two bonded atoms j and k was examined. In idealized α-Te, the bonding angle θ_{jik} (where atom i is the central atom) should be around 103°. Given thermal fluctuations of atomic environments at finite temperatures, a threshold was manually picked up which was used to rule out too large or too small values of θ_{jik}: namely, only atoms with $\theta_{jik} \in [80°, 160°]$ were considered as chiral fragments (Fig. 4.12a).

Lastly, the possible helix, in which a given atom would take part, was examined. Given the connectivity of three atoms that satisfies the criteria in the first and second steps (viz. atoms i, j, and k, in which atom i is bonded to j and k), the atomic chirality is still ambiguous (cf. Fig. 4.12). Hence, a fourth atom l, which is bonded with either atom j or atom k, is needed to fully define the per-atom chirality. Notably, a circumstance was also ruled out when these four atoms are nearly in the same plane, which is based on the calculation of the dihedral angles of these four atoms. The vectors $\vec{v}_{ij}$ (with a direction from atom i to atom j) and $\vec{v}_{ik}$ were computed (with a direction from atom i to atom k), and therefore the normal vector of the ijk plane can be obtained as the cross product

$$\vec{v}_{\text{norm}} = \vec{v}_{ij} \times \vec{v}_{ik}.$$

Next, the connectivity of the fourth atom was checked, l. On one hand, if atom l is connected to atom j, another vector $\vec{v}_{lk}$ (with a direction from atom l to atom k) can be obtained which is parallel to the chain direction of this helical fragment. The dot product of $\vec{v}_{\text{norm}}$ and $\vec{v}_{lk}$ was then computed as

$$c = \vec{v}_{\text{norm}} \cdot \vec{v}_{lk}.$$

If $c > 0$, then the fragment is left-handed; and if $c < 0$, then the fragment is right-handed (Fig. 4.12). On the other hand, if atom l is connected to atom k, a vector $\vec{v}_{lj}$ can also be computed (with a direction from atom l to atom j). The dot product of norm and $\vec{v}_{lj}$ was then given as

$$c = \vec{v}_{\text{norm}} \cdot \vec{v}_{lj}.$$

If $c > 0$, then the fragment is right-handed, and if $c < 0$, then the fragment is left-handed (Fig. 4.12).

Notably, the thermal fluctuation of atoms at finite temperatures results in considerable structural noise (e.g., local bond distortions). Hence, to better define the chirality of different chiral grains, the structural model was relaxed using the ACE potential to remove the effect of thermal fluctuation.

To compute the number of grains in MD simulations, a graph clustering algorithm described in Ref. 330 was employed to separate chiral atoms with different chain orientations (i.e., grain segmentation). A graph was constructed based on any given atomic configuration: atoms were taken as nodes and chemical bonds were taken as edges. The edge weights of the graph were defined as: $W = e^{-\theta^2/3}$ (as also done in Ref. 258), in which θ is the angle between the chain direction (i.e., $\vec{v}_{lk}$ or $\vec{v}_{lj}$ in Fig. 4.12) and the c-axis (i.e., the [001] direction). The grain segmentation was only performed on left- and right-handed grains, but separately. In other words, grains with a different chiral shape but the same chain direction will be regarded as two different (chiral) grains.

The per-atom chirality can thus be quantified, using the aforementioned measure based on local geometry and helical connectivity, which is used to classify atoms as left-handed, right-handed, or achiral. A random distribution of left- (blue) and right-handed (red) chiral grains was found. Two chiral grains with different chain

directions were connected by a disordered boundary (Fig. 4.9a). Averaged over five independent samples, the numbers of left- and right-handed chiral-environment atoms during crystallisation are very similar (Fig. 4.9e), consistent with the fact that the routes from the hypothetical cubic phase to left- and right-handed helices are energetically equivalent (Fig. 4.9).

4.5 Boundary ordering and chirality transfer

Looking beyond the formation of chiral chains, excess kinetic energy at higher temperatures likely results in direct chirality transfer between left- and right-handed chains (Fig. 4.4b). To reveal the mechanism of such chirality transfer, the crystallised structural model at 323 K was further annealed at higher temperatures using ACE-MD simulations, from 323 to 473 K in an increment of 50 K. The post-annealing processes were shown in Fig. 4.13.

In particular, Fig. 4.14a shows marked grain-boundary ordering and chirality transfer in polycrystalline Te at 473 K. Early on during annealing, small chiral grains were consumed by adjacent larger grains. The surviving chiral grains continuously competed with each other, leading to frequent chirality transfer and grain-boundary ordering.

According to classical nucleation theory, the annihilation of small unstable grains stems from their large interfacial energies compared to the bulk. Hence, these grain-boundary ordering and chirality transfer were driven by the formation of energetically more favourable interfaces between neighbouring grains, as evidenced by the ML per-atom energies (Fig. 4.15).

Unlike DFT, which predicts the total energy of a given structure, ML potentials predict per-atom energies [162, 164], and the total energy of a structure is the sum of the energies of each atom. As such, these local energies can be probed to investigate local energetics in materials. GAP local energies have been used to provide insightful energetic information in many theoretical works, such as the energy distributions of disordered graphene [331], coordination defects in a-Si [332], and the

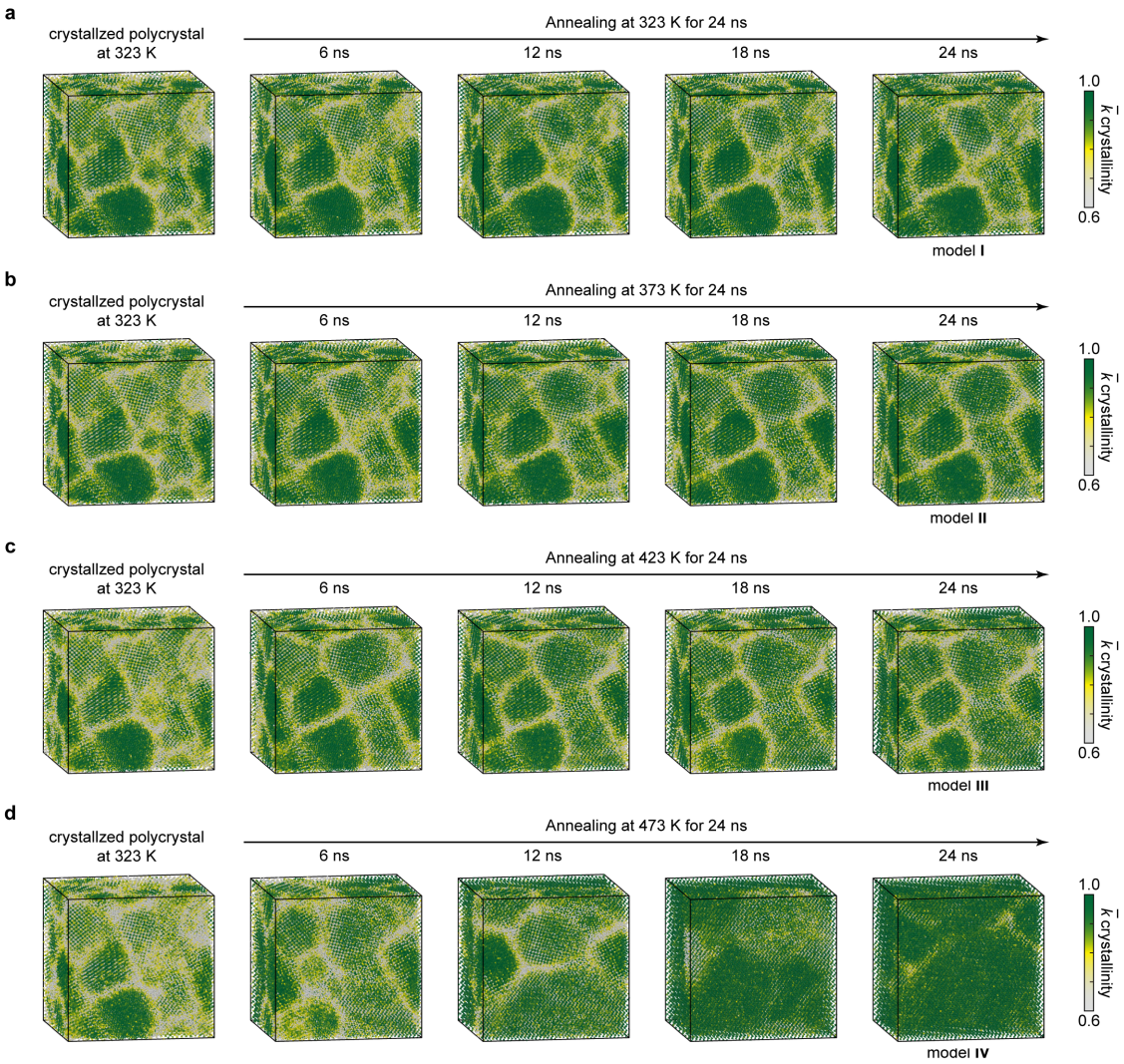


Figure 4.13: The trajectories of post-annealing simulations. The recrystallised polycrystalline Te sample (as shown in Fig. 4.9a) was annealed at various temperatures for another 24 ns, including (a) 323 K, (b) 373 K, (c) 423 K, and (d) 473 K. Atoms are colour-coded using the per-atom SOAP crystallinity $\overline{k}$, illustrating a gradual structural ordering process.

local stabilization of simple-hexagonal (sh) crystallites during the pressurization of amorphous Si [279]. Given this successful track record of using GAP local energies to study local structure, the developed GAP model was used to compute the total energy and per-atom energies of the structural models.

The atomistic routes of chirality transfer were then investigated, and two different types were found (Fig. 4.14c–d). The first type occurred at the boundary between two chiral grains with different chain directions, proceeding via gradual atomic rearrangements at the interfaces. Figure 4.14c shows a representative tran-

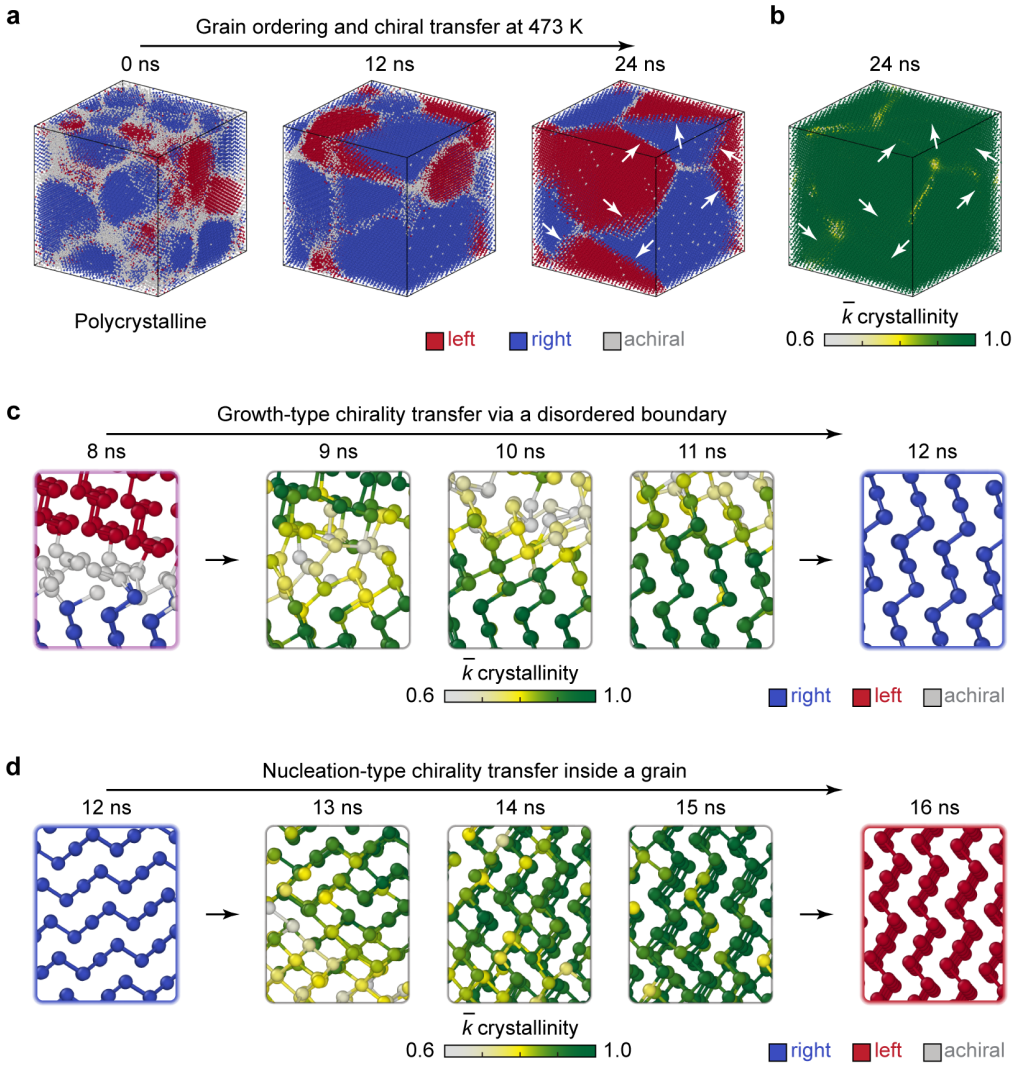


Figure 4.14: Grain-boundary ordering and chirality transfer in Te. (a) Annealing the crystallised Te model (Fig. 4.9a) at 473 K for another 24 ns. Atoms are colour-coded according to their per-atom chirality. White arrows indicate the sharp boundaries between the grains. (b) The fully annealed structure, colour-coded using the $\overline{k}$ crystallinity measure, in the same way as shown in Fig. 4.9a. White arrows indicate the boundary between two chiral grains. (c, d) Typical snapshots showcasing two different types of chirality transfer during the annealing process: (c) growth-type and (d) nucleation-type chirality transfer. Atoms are colour-coded using per-atom SOAP crystallinity $\overline{k}$ or per-atom chirality.

sition process: many amorphous-like atoms ($\overline{k} < 0.75$) were found at a disordered boundary between two adjacent grains. Given excess kinetic energy, these atoms rearranged and formed energetically more favourable cubic-like motifs. A marked similarity exists between this chirality transfer and the growth process (Fig. 4.11) during the crystallisation process. Both quickly proceeded at the interface, via

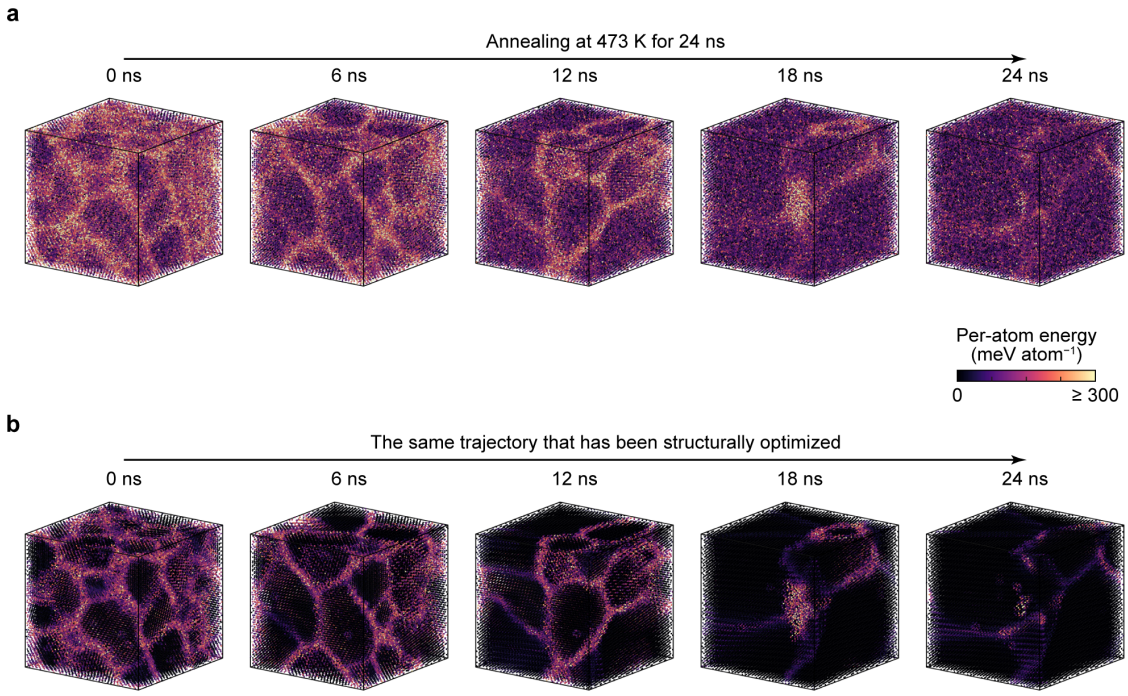


Figure 4.15: The trajectories of high-temperature annealing simulations at 473 K, with atoms colour-coded to indicate the per-atom energy (in a unit of meV atom^{-1}). The structural models were directly taken from the MD trajectories at 473 K (a) without structural optimisation, and (b) with structural optimisation. The reference energy (= 0 meV atom^{-1}) is taken from the energy of the optimised α-Te.

formation of disordered cubic-like motifs, leading to chiral fragments. The grain-boundary ordering was mostly dominated by this growth-type chirality transfer, rapidly reducing the number of grains (Fig. 4.16).

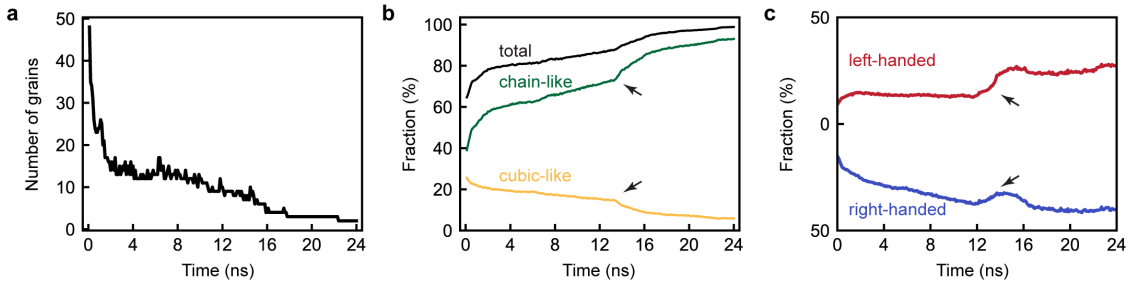


Figure 4.16: Quantitative structural analysis for high-temperature annealing, including (a) the calculated number of grains; (b) the counts of crystal-like (i.e., cubic-like and chain-like) atoms; and (c) the calculated fraction of Te atoms in left- and right-handed chiral helices during the annealing process. Black arrows in (b) and (c) indicate multiple nucleation-type chirality transfers upon annealing.

The second type of chirality transfer proceeded in a more collective way. Due to considerable thermal fluctuations, chiral atoms inside crystal grains might temporarily become achiral but remain in cubic-like motifs (cf. Fig. 4.4). In analogy to the collective structural transition to chiral helices in nucleation, such cubic-like fragments were not stable, and they either returned to the original chiral structures or reformed other chiral helices with different handedness or chain directions (Fig. 4.14d). A remarkable appearance of this nucleation-type chirality transfer was found at 13–14 ns, after two large grains connected via sharp boundaries, resulting in a sudden change in the left-to-right-handed ratio (Fig. 4.16). In the MLMD simulations, all disordered grain boundaries were eliminated at the end of reordering, and no chirality transfer was evident in the stable chiral grains.

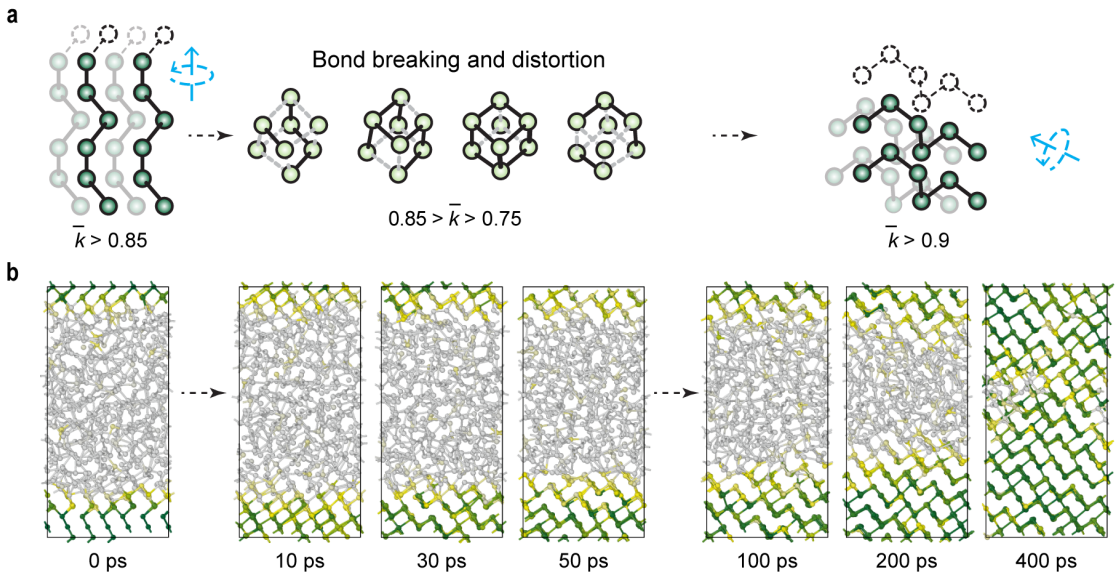


Figure 4.17: (a) A sketch depicting a change of chain directions via the formation of disordered cubic-like motifs in the growth simulations. (b) Growth simulation along the [001] direction; however, the chain direction changed during the growth. Atoms are colour coded using the per-atom SOAP crystallinity $\bar{k}$, illustrating a gradual structural ordering process.

In fact, the change in chain direction was also observed in small-scale growth simulation with a prefixed crystalline template using ab initio molecular dynamics (Fig. 4.17), which was also mediated by the formation of cubic-like motifs. Starting from the initial growth direction of [001], cubic-like motifs formed quickly at the

starting of growth. These cubic-like motifs suddenly and collectively turned into another chiral fragment, via frequent bond breaking and reformation, which adopted a different chain direction but the same chiral shape. The growth continued after the change of chain direction, leading to a recrystallised structural model but with a different chiral direction as compared to the initial one.

After the deep annealing at 473 K, only two chiral grains, one left- and one right-handed, were found. These two grains had the same chain direction, nearly perpendicular to the z-axis of the simulation box, and were separated by sharp boundaries. The well-annealed structural model was colour-coded by per-atom chirality and $\overline{k}$ crystallinity (Fig. 4.14b), showing that the chiral grain boundaries are highly crystal-like. Moreover, the computed per-atom energy of the fully annealed structure also showed the energetically stable grain boundaries (Fig. 4.18).

In addition, both chiral grains show pyramid-like morphology with sharp facets (Fig. 4.19). These observations from the atomistic simulations using the ACE potential are consistent with the 3D chiral morphology characterized using STEM [291], in which the synthesized Te nanoparticles showed pyramid-like shapes. Moreover, screw dislocations were observed at the boundary of one chiral grain (Fig. 4.18), in agreement with atomic-resolution STEM [291], which was believed to mediate the formation of morphographic chirality by creating a more reactive crystal-growth front [333–336]. Taken together, this evidence implies an atomic-level origin of morphographic chirality—likely due to forming screw dislocations and stable facets during growth and ordering—instead of resulting from crystallographic chirality alone.

4.6 Switching in Te selector devices

As mentioned in Section 4.1, due to the volatile and non-linear electric response in phase-change processes, Te has been regarded as a promising candidate for selector devices, which has been reported in Ref. 301. This section explores the microscopic switching mechanism of Te selector devices [301]. The structural model used here measures about $15 \times 15 \times 15$ nm^3, comparable to real devices with 10–20-nm thick

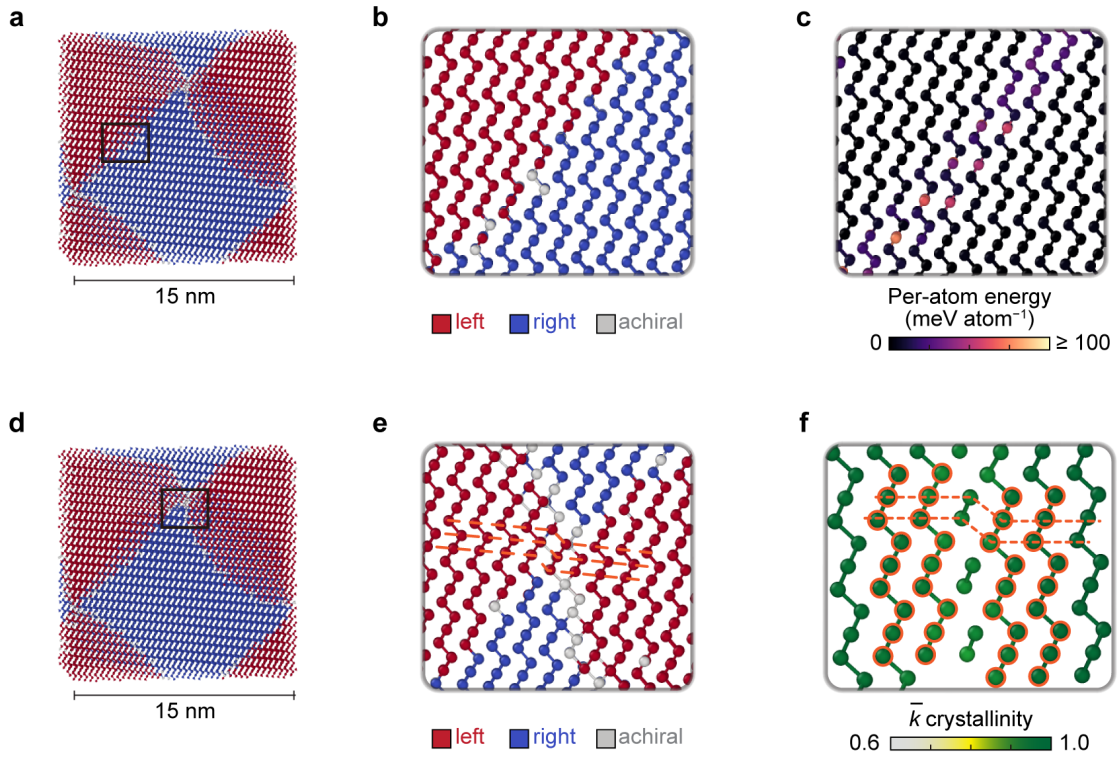


Figure 4.18: Coexistence of regions with two chiral chains and the formation of screw dislocations in the well-annealed structure. (a, d) The same slice of the well-annealed structural model. Atoms are colour coded using per-atom chirality. The black triangles mark two areas where left helices were directly connected with right helices. (b, e) Magnified view of the regions marked in (a) and (d), respectively. (c) The same view as in (b) but with atoms now labelled using the per-atom energy. The reference energy (= 0 meV atom^{-1}) is taken from the energy of optimized α-Te. (f) Magnified view based on panel (e) where atoms are labelled using the per-atom SOAP crystallinity, $\overline{k}$. Orange dashed lines in (e) and (f) emphasise the trend of forming screw dislocations.

Te films [301]. During manufacturing, devices are kept at high temperatures (e.g., 400–450 °C) for hours [25]. As shown by the simulations in Fig. 4.14, such high-T annealing can lead to significantly fewer and larger grains. Hence, the initial state of Te films in the selector device is likely to be a single crystal. Figure 4.20a shows typical experimental current–voltage data of selector devices in the first switching cycle [300], exhibiting volatile switching behaviour [2, 24]. As the external voltage increases to the firing voltage (V_{fire}), the device first switches to the low-resistance ON state. To model this process, a 5-ns heating simulation was performed on a 102,000-atom supercell of α-Te, representing the ideal case of starting from a

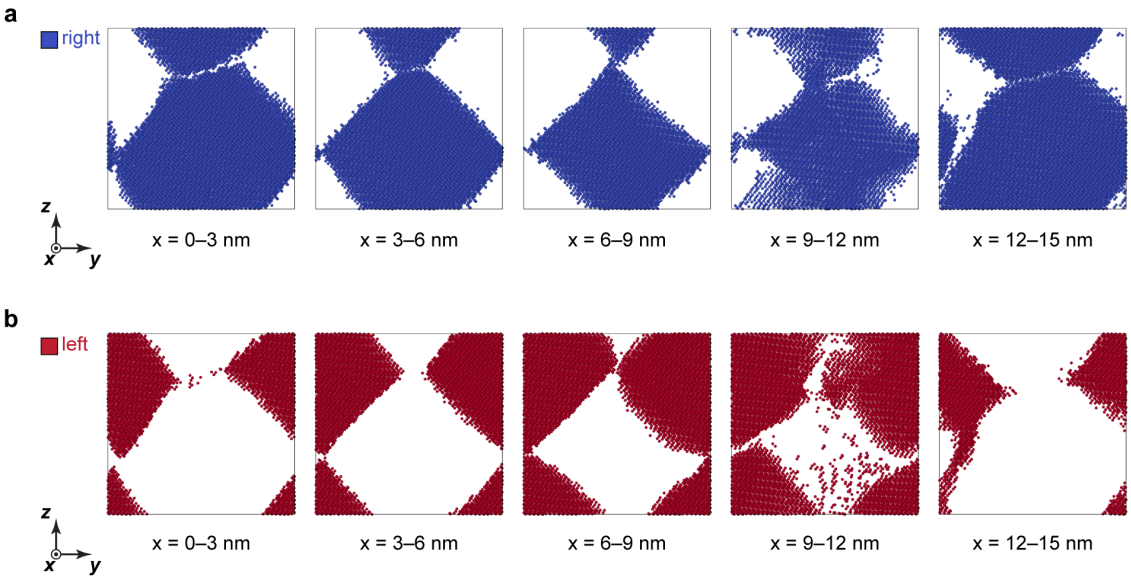


Figure 4.19: Morphography of two chiral grains in the well-annealed structural model upon high-temperature annealing. (a) Different parallel slabs (with a thickness of 3 nm) perpendicular to the x-axis. Only atoms in right-handed helices are shown. (b) Different parallel slabs (with a thickness of 3 nm) perpendicular to the x-axis. Only atoms in left-handed helices are shown. Considering the periodic simulation box, both chiral grains show pyramid-like morphology with sharp facets.

perfect single crystal (Fig. 4.20b). The structure did not melt until reaching the experimental melting temperature of 723 K (Fig. 4.20c).

In experiments, the selector can be switched off within ≈13 ns, after reducing the voltage below the holding voltage (V_{hold}) [301]. In the second switch, only a small threshold voltage (V_{th}) was required to set the ON state (Fig. 4.20d). To understand such different switch-on voltages, the melted structure from the first cycle was cooled down from 800 to 300 K within 5 ns. The structural model was kept at 300 K for another 12 ns until it fully crystallised (Fig. 4.21). This crystallisation mainly occurred at low temperatures (below 350 K), initialised by fast nucleation via the fast formation of cubic-like motifs. The whole process was consistent with the observations of the crystallisation shown in Fig. 4.9a, resulting in polycrystalline Te. Next, this model was heated up for another 5 ns (Fig. 4.20e). Different from the melting process of single-crystalline α-Te (Fig. 4.9b), the polycrystalline structure started to collapse via the grain boundaries at 523 K (Fig. 4.20f)—much lower

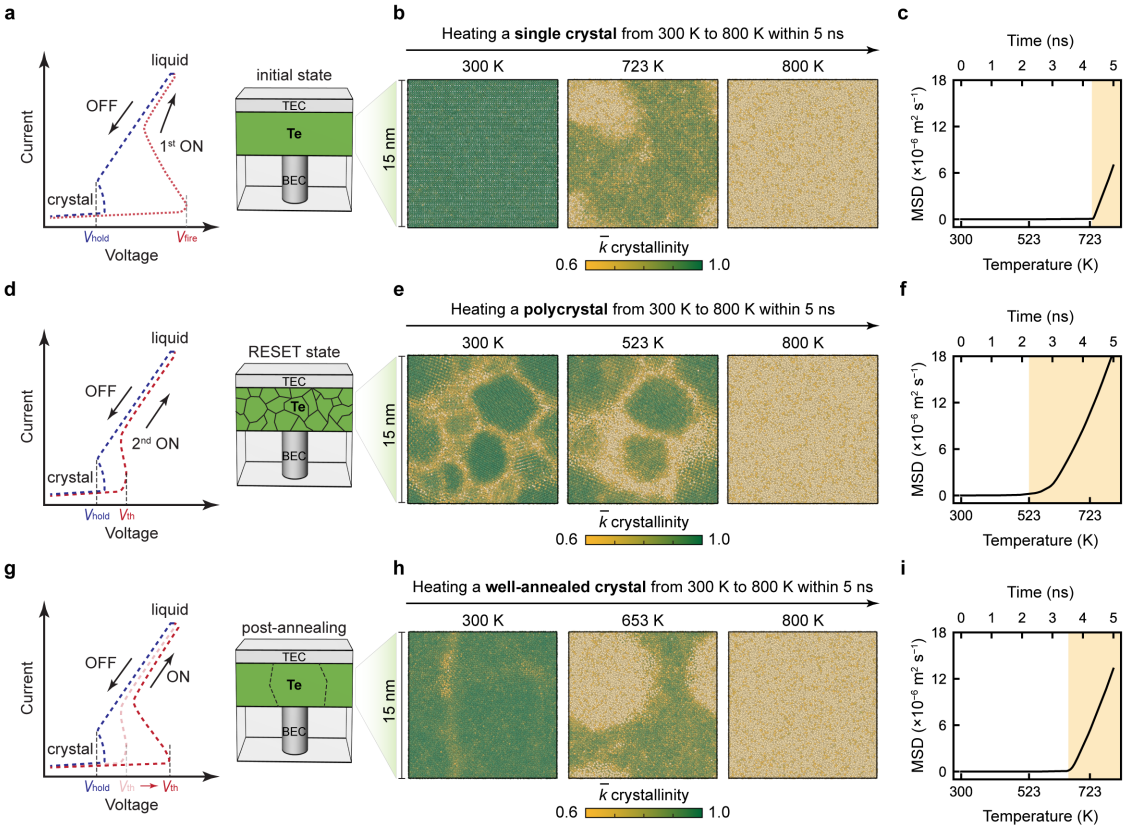


Figure 4.20: Device-scale switching process of Te. (a) A schematic experimental current–voltage curve during the first switch of Te selectors, in which the initial state is a single crystal. Inset: sketch of Te-based selectors in memory devices reported in Ref. 301. The Te film, with a thickness of 10 to 20 nm, is sandwiched by top and bottom electrode contacts (TEC and BEC). (b) A 5-ns simulated heating process (from 300 to 800 K) of single-crystalline Te. Atoms are colour-coded by per-atom SOAP crystallinity $\overline{k}$, similar to Fig. 4.9a. (c) The calculated mean-square distance (MSD) of atoms during the heating process. (d) A schematic experimental current–voltage curve during the second switch of Te selectors, in which the Te film is polycrystalline. (e, f) As panel (b, c), but now for a polycrystalline sample, showing distinctly faster melting. (g) A schematic experimental current–voltage curve demonstrating the increased switch-on voltage upon thermal annealing at high temperatures. (h, i) As panel (b, c), but now for a post-annealed sample with few chiral-grain boundaries, showing an increased melting point as compared to the polycrystal structure.

than the normal melting temperature of Te (723 K). Hence, given abundant grain boundaries as melting "seeds" in the polycrystal after nanosecond crystallisation, the V_{th} required for the second switch is expected to be smaller than the V_{fire} needed to melt a single crystal.

In fact, considerable noise occurred for the switch-on voltages (V_{th} and V_{fire}) in Te selectors [301], which was suggested to result from different degrees of polycrys-

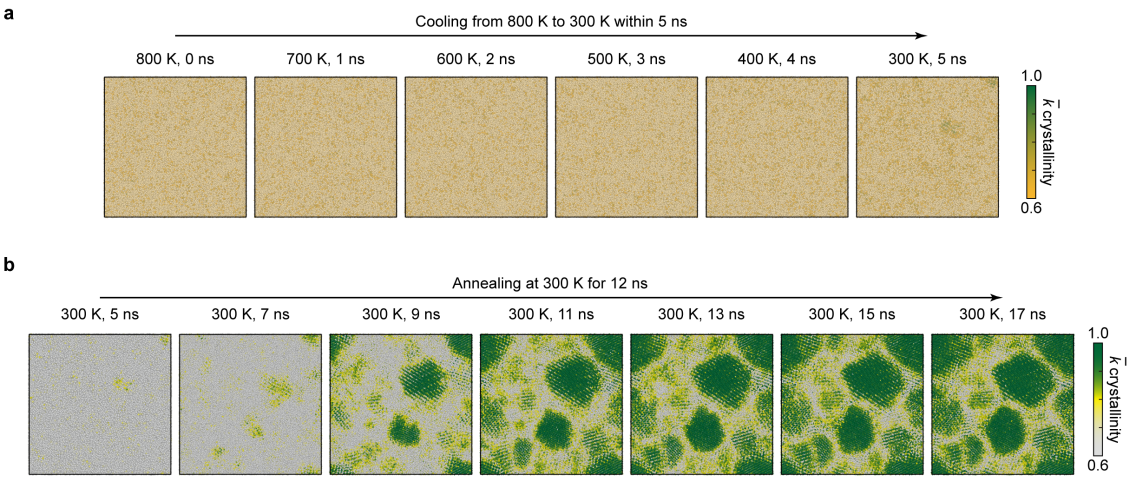


Figure 4.21: The crystallisation of Te during the cooling process. (a) The 5-ns cooling process (from 800 to 300 K) of liquid Te. (b) The 17-ns crystallisation process of Te at 300 K. Atoms are colour-coded using two different colour schemes to represent the per-atom crystallinity, which separate the cooling process in (a) and the crystallisation process (when annealing at 323 K) in (b).

tallinity in the recrystallized material [337]. To understand this, the five independent crystallisation simulations at 323 K (cf. Fig. 4.8) were examined, and different grain topographies were found in the recrystallized structures (Fig. 4.8). Such structural stochasticity from the nanosecond crystallisation contributes to the noise in switch-on voltages. Moreover, an increased switch-on voltage was observed in Te selectors upon thermal treatment at elevated temperatures from 100 to 400 ℃ [301]. To understand this, the post-annealed structures (i.e., the last snapshots from different high-temperature annealing trajectories shown in Fig. 4.13) were taken as the starting configurations for another 5-ns heating simulations (Fig. 4.22). Different melting behaviours were observed in different structural models.

When heating the structure that had been post-annealed at 473 K, the structural-disordering collapse only occurred at the sharp boundaries between two chiral grains at 653 K (Fig. 4.20h), and expanded across the entire structural model. In contrast, when heating the post-annealed structure that have more grain boundaries at lower temperatures, the starting of melting behaviours was observed at the grain boundaries, resulting an earlier melting as compared to the post-annealed structure at 473

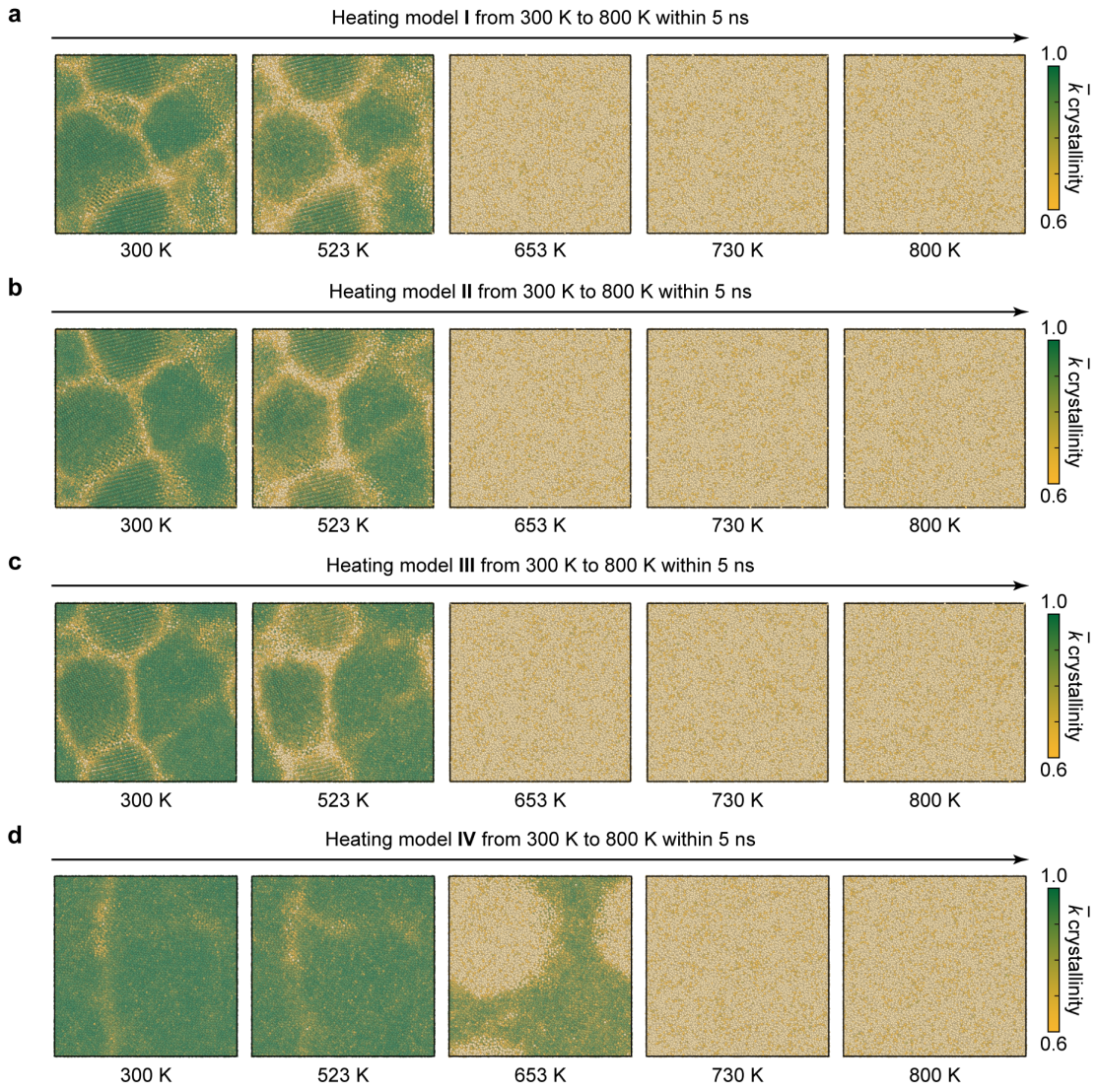


Figure 4.22: The heating processes of different polycrystalline samples. After the recrystallisation of Te (as shown in Fig. 4.9a) at 323 K, the polycrystalline structure was then post-annealed at four different temperatures, viz. 323, 373, 423, and 473 K, for another 24 ns (cf. Fig. 4.13). The last configuration of these four post-annealing trajectories was taken, labelled as model **I** to **IV**. Four 5-ns heating simulations (from 300 to 800 K) were performed on (a) model **I**, (b) model **II**, (c) model **III**, and (d) model **IV**. Atoms are colour coded using the per-atom SOAP crystal-similarity $\overline{k}$, illustrating a gradual loss of structural ordering upon heating.

K. Hence, the shifted switch-on voltage is a direct consequence of grain numbers and sizes in the OFF-state structure. Post-annealing the structure at higher temperatures results in fewer grain boundaries and less structural stochasticity, leading to much more similar melting behaviours and thus giving more consistent switch-on operations in the Te selector devices. The increased melting point (e.g., via post-

annealing pulses) also implies a larger switch-on voltage, which is critical for selector devices to provide higher operation windows [45] and alleviates operation disturbance issues [338] for crossbar memory devices. Hence, additional post-annealing pulses could be applied to reduce the structural stochasticity of the OFF-state before the subsequent switch-on operations, which could simultaneously improve the programming consistency and enlarge the V_{th} value for practical use.

4.7 Conclusion

In this chapter, machine-learned interatomic potentials were developed, based on an iterative training process (as mentioned in Section 3.2), and then used in a comprehensive study on the crystallisation and chirality transfer in elemental tellurium. Based on quantum-accurate and realistic-scale ML-driven simulations (compatible to the size of real-world devices), the microscopic mechanisms of crystallisation and chirality formation in elemental Te, the simplest chiral inorganic system, have been revealed. Instead of directly forming helical chains, the rapid nucleation in Te starts with higher-coordinated, disordered cubic-like seeds. With only nanosecond lifetimes, these transient motifs undergo fast and collective distortions to form chiral helices. When annealing at high temperatures, e.g., 473 K, polycrystal Te undergoes an evident chirality transfer process. Two different types of chirality transfer were found, resulting in a marked grain-boundary-ordering process, which were also mediated by disordered cubic-like motifs.

Moreover, the large-scale MLMD simulations, with a similar size of real-world devices and over tens of nanoseconds, allowed to explain the switching mechanisms of Te-based selector devices on a fundamental level. The volatile switching behaviour of Te-based devices results from a large crystallisation tendency of Te, showing swift and spontaneous crystallisation upon cooling from high temperatures. Such large tendency is supported by the formation of cubic-like motifs which largely reduced the stochastic nature of nucleation process. It is also found that the degree of polycrystallinity is key to determining the switching voltage: the fewer grain numbers

(or larger grain sizes) there are in the OFF-state structure, the larger the switch voltage. Given that high-temperature thermal annealing will result in evident grain-boundary ordering process, increasing the sizes of grains (or decreasing the number of grains), via post-annealing, will help reduce the programming stochasticity and thus increase the consistency of Te-based devices.

Notably, this chapter has provided a good example of how ML potentials can be developed and used in the study to solve textbook questions in structural chemistry (e.g., the chirality transfer in tellurium) and also to explore switching mechanisms in practical Te-based selector devices. In addition, this chapter has revealed the atomic-level origins of chirality in α-Te and can provide a starting point for realistic computational studies of other chiral crystals.

Chapter 5

ML-driven modelling of Ge–Sb–Te phase-change materials

This chapter discusses the development of an ML potential for Ge–Sb–Te phase-change materials, which enables large-scale MD simulations over a long time scale. Note that part of this section is based on a previously published work in Nat. Electron. 2023 (Ref. 180). Some texts and figures have been reused. I would also like to thank En Ma, Wei Zhang, and Volker L. Deringer for their guidance and useful discussions in this work.

5.1 Atomistic modelling of Ge–Sb–Te alloys

To start with, a brief recap is provided about phase-change materials (PCMs) and atomistic modelling on PCMs, which have all been introduced in Chapter 1. PCMs is a leading material candidate for non-volatile memory and neuro-inspired computing technologies [1, 25, 339, 340]. PCMs can be switched reversibly and rapidly between two solid-state phases, via crystallisation (SET) and amorphisation (RESET), respectively. A large property contrast between the amorphous and crystalline phases, both in electrical conductivity and optical reflectivity, is used to encode digital information. The most useful PCMs are located on the quasi-binary GeTe–Sb_2Te_3 line of chemical compositions ("GST"). For example, $Ge_2Sb_2Te_5$ and $Ge_8Sb_2Te_{11}$ are used in re-writeable optical disks [339], $Ge_1Sb_2Te_4$ has been claimed as the key ingredient for the 3D cross-point memory products [8], and new Sb_2Te_3-based PCMs with suitable dopants or nanoconfinement layers are promising candidates for universal memory [110, 341] or neuro-inspired computing devices [47]. When integrated with waveguides, $Ge_2Sb_2Te_5$ enables various on-chip non-volatile photonic applications, including memory, computing, colour rendering, and nanopixel displays [342].

Computer simulations based on DFT and DFT-driven AIMD have provided key insights into structural features [89, 343, 344], bonding contrast [345, 346], and crystallisation kinetics of GST [101, 105, 109, 114]. However, the computational cost of AIMD prevents simulations of systems beyond a thousand atoms or so, and of processes taking more than a few nanoseconds [114]; at the same time, the structural and chemical complexity in PCMs calls for quantum-mechanically accurate methods rather than empirically parametrised force-field models. These requirements have made it challenging, and often impossible, to simulate more realistic scenarios encountered in devices: Joule heating induces a temperate gradient by programming pulses; the active volume expands and shrinks during switching; the local chemical composition evolves; and model system sizes of tens of nanometres are required to account for the "real-world" device geometry.

ML based interatomic potentials are an emerging approach in the field to address the challenges mentioned above, aiming to combine the efficiency of empirical potentials and the accuracy of DFT [118]. Based on a reference database of small-scale DFT computations, the potential energy surface becomes the regression target for an ML model, allowing for fast and accurate simulations. As discussed in Section 1.5, many ML potentials have been developed for different PCMs in the past decade, e.g., GeTe [119], $Ge_2Sb_2Te_5$ [135, 347], Sb [132], etc. However, thus far, these models have been designed and verified for specific compositions, and a comprehensive model for the entire GeTe–Sb_2Te_3 quasi-binary line has remained a challenge.

In this chapter, a qualitative step forward in the atomistic modelling of PCMs is discussed, demonstrating that fully atomistic device-scale simulations are made possible by ML potentials. It is shown that ML-driven modelling can address questions that are key in novel PCM architectures, such as the simulation of cumulative SET and iterative RESET processes—thereby revealing the atomistic processes and mechanisms in PCMs on the full ten-nanometre length scale that is relevant to devices.

5.2 Fitting ML potentials for Ge–Sb–Te PCMs

The quality of any ML model hinges on the quality of its input data, and the performance of an ML potential depends in no small part on the quality of the reference database to which it is fitted [171]. Here, a new dataset of representative structural models (data locations) and DFT-computed energies and forces (data labels) was introduced as a platform for fitting ML potentials for GST PCMs. Starting with an initial set including liquid and amorphous configurations from AIMD, structural space was iteratively explored with evolving ML potentials and progressively trained the models to represent more complex scenarios (Fig. 5.1a). Iterative training allows for efficient reference data generation [171], because no further AIMD is required beyond the initial set of data, which was called "iter-0". It is followed by performing GAP-driven melt–quench simulations, of the type that are common in the field ("iter-1"), and then followed with more specialised simulations, designed to reflect the particular requirements of the do-main of application ("iter-2"). The latter allows the model to "learn" microscopic information most relevant to GST, i.e., about melting and crystallisation processes in which phases with different degrees of structural order coexist. The "iter-2" part is central here, because it gradually explores and includes the domain knowledge that is relevant to how the PCM functions and evolves in a real-world action.

5.2.1 Reference database based on iterative training

The main potential, referred to as "GST-GAP-22" in the following, is intended to serve as a general-purpose model that can be used "off the shelf", without further modification by the user. In this, the approach is complementary to the recently reported "on-the-fly" fitting of potentials during an ab initio molecular dynamics (AIMD) simulation [348]. The latter provides efficient sampling of small-scale reference data, however requiring DFT to "kick-start" the process initially. Hence, the on-the-fly ML potentials can be viewed as highly promising for future work for a

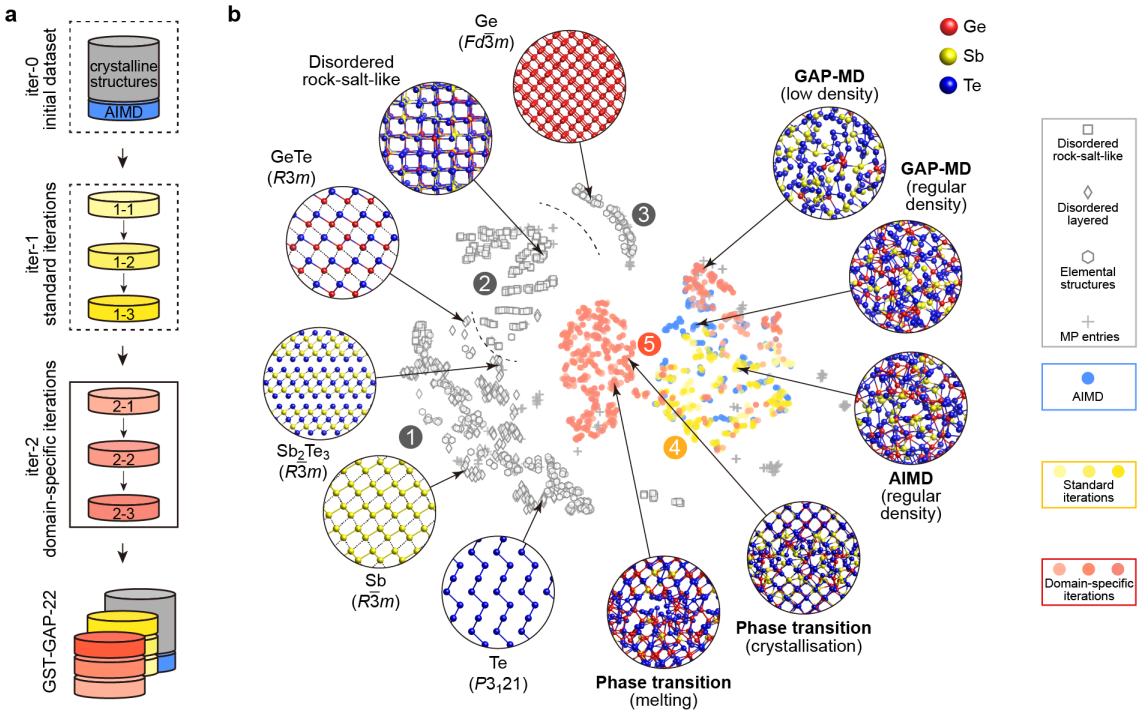


Figure 5.1: Machine-learned potential fitting for Ge–Sb–Te phase-change materials. (a) Reference data generation for increasingly comprehensive GAP ML models. The initial reference data-base ("iter-0") contains crystalline structures and ab initio MD (AIMD) snapshots of liquid and amorphous phases. A two-step iterative training process is then carried out, as described in the text, with standard iterations ("iter-1") and domain-specific iterations ("iter-2") gradually expanding the database. (b) A two-dimensional structural map representing the reference database, generated using the SOAP similarity matrix. Representative structures are shown, and are discussed in the text according to the labels (1)–(5).

specific range of configurations in a relatively small area of structural space. The present work is conceptually different, which relies on extensive domain knowledge to build a generally applicable fitting database and ML model that does not require any new DFT data at runtime.

The reference database contains several different parts which are described below; and they are summarised in Table. 5.1.

Free atoms and dimer data

All three isolated atoms (Ge, Sb, and Te) were included, each in a $20 \times 20 \times 20$ Å^3 cubic simulation cell. These cells provide energy references for the fit. The dimer configurations were considered for all six different atomic pairs, viz. Ge–Ge, Ge–Sb,

Table 5.1: Summary of the composition of the GST-GAP-22 reference database.

		Database size		Sparse
		Cells	Atoms	points
Free atoms	Ge / Sb / Te	3	3	1
Dimer data	Ge–Ge/Ge–Sb/Ge–Te/Sb–Sb/Sb–Te/Te–Te	210	420	90
iter0	crystalline structures	1261	69055	90
	AIMD data	210	41490	1200
iter1-1		112	21994	600
iter1-2	Disordered (commonly relevant densities)	112	21994	600
iter1-3		112	21994	600
iter2-1	Crystallisation configurations	84	21876	
	Melting configurations	84	21770	1200
	Disordered (more open structures)	56	10992	
iter2-2	Crystallisation configurations	84	21876	
	Melting configurations	84	21914	1200
	Disordered (more open structures)	56	10992	
iter2-3	Crystallisation configurations	84	21876	
	Melting configurations	84	21914	1200
	Disordered (more open structures)	56	10992	
Total		**2692**	**341132**	**8191**

Ge–Te, Sb–Sb, Sb–Te, and Te–Te, again described using 20 × 20 × 20 Å^3 cubic simulation cells. For these dimer scans, various interatomic distances were sampled between 1.0 and 7.0 Å, in which an interval of 0.1 Å was used from 1.0 to 2.0 Å and an interval of 0.2 Å from 2.0 to 7.0 Å.

The initial "iter-0" dataset

The aim of this dataset is to facilitate the preliminary exploration of different structural phases of GST and to achieve a reasonable description of crystalline and disordered GST for the subsequent iterative rounds of training. The dataset consists of two parts, namely, crystalline GST structures and disordered configurations from AIMD trajectories.

In terms of crystalline phases, the ground-state structures of the constituent elements were considered, viz. Ge (diamond-type, $Fd\bar{3}m$), Sb (layered, $R\bar{3}m$), and Te (helices, $P3_121$). For seven GST compositions, both primitive cells and supercells of the stable (trigonal) phases were included, viz. Sb_2Te_3 ($R\bar{3}m$), $Ge_1Sb_4Te_7$ ($P\bar{3}m1$),

$Ge_1Sb_2Te_4$ ($R\bar{3}m$), $Ge_2Sb_2Te_5$ ($P\bar{3}m1$), $Ge_3Sb_2Te_6$ ($R\bar{3}m$), and GeTe ($R3m$), and their metastable (rock-salt-like) phases in $2 \times 2 \times 2$ or $3 \times 3 \times 3$ supercell expansions, with atomic vacancies randomly distributed on the cation sites. In addition to the ideal crystalline forms at varied (scaled) lattice parameters, various copies of the structural models with applied distortions were also added to better sample the potential-energy surface (PES) of crystalline GST. Some cation-like atoms (Ge and Sb atoms) were manually swapped with the anion-like ones (Te atoms) to sample anti-site defects. The crystalline structures obtained from the Materials Project (MP) database were also prepared [320], which included both the stable phases and certain metastable structures, with the goal to sample a more diverse range of configurations.

The disordered configurations in the "iter0" dataset also contain seven GST compositions, viz. Sb_2Te_3, $Ge_1Sb_4Te_7$, $Ge_1Sb_2Te_4$, $Ge_2Sb_2Te_5$, $Ge_3Sb_2Te_6$, $Ge_8Sb_2Te_{11}$, and GeTe, in cubic simulation cells containing fewer than 200 atoms, which were taken from AIMD melt-quench trajectories at different temperatures. Varied lattice parameters for these configurations were considered, ranging from 17.7 to 19.2 Å with an interval of 0.3 Å, corresponding to commonly relevant atomic densities for frequently visited scenarios, that is, the highest density was chosen slightly larger than that of crystalline rock-salt-like GST, whereas the lowest is slightly smaller than the theoretical density of the amorphous phase at the PBE level [147].

Standard iterations (iter1-1, iter1-2, and iter1-3).

The datasets obtained from standard iterations only include disordered configurations obtained from iterative GAP-driven melt-quench simulations. The objective of these three datasets is to extend the sampling of disordered configurations and thereby to enable the potential to accurately describe the structural properties of GST. I considered the same seven compositions as above (Sb_2Te_3, $Ge_1Sb_4Te_7$, $Ge_1Sb_2Te_4$, $Ge_2Sb_2Te_5$, $Ge_3Sb_2Te_6$, $Ge_8Sb_2Te_{11}$, and GeTe), and the various densities in the same range as in iter0. The models were generated following the same

melt-quench protocol.

Domain-specific iterations (iter2-1, iter2-2, and iter2-3).

Each "iter2" dataset consists of three different kinds of configurations: corresponding to crystallisation, to melting, and to structures at even smaller (than the widely used ones) atomic densities. All three datasets include the same seven stoichiometric compositions as those in iter0 and iter1. The respective computational tasks are chosen so as to mimic larger-scale simulations of crystallisation and melting as far as possible within the system-size limits of DFT.

To generate intermediate configurations for the crystallisation process, the starting configuration were rock-salt-like GST structures (with manually created anti-site defects as discussed for the "iter0" dataset) and two atomic layers were fixed to represent the crystalline matrix during the melt–quench process. The structural models were then annealed at 600 K without any geometric constraints and gradually crystallised, forming crystalline–amorphous interfaces which are represented by this part of the "iter2" dataset.

To obtain the melting configurations, the initial structures (i.e., both rock-salt-like and trigonal GST) were heated up from 300 K to 1,200 K without any geometric constraints. Intermediate configurations were taken from the heating process, in which different scales of atomic distortions were observed. Density variations were also considered, including commonly relevant values.

The generation of configurations at even smaller atomic densities (than the commonly relevant ones) follows the same melt-quench protocol as discussed above. These configurations were included to ensure the potential is stable at low densities as might occur, e.g., during volume fluctuations in a constant-pressure high-temperature simulation.

The composition of the reference database is summarised and illustrated in Fig. 5.1b. In this two-dimensional structural map, the distance between any two points indicates the SOAP kernel similarity [170] of the corresponding configurations. The

areas labelled (1)–(3) represent crystalline structures included from the start (grey points): the first two comprise ordered and disordered GST, respectively, for which the key structural motif is an (often defective) octahedral environment; these points are distinct from area (3), representing diamond-type Ge. Area (4) covers liquid and amorphous GST configurations, mostly added during early iterations (yellow points); area (5) reflects phase-transition snapshots, added during the domain-specific series (red points), and the points are located between the fully ordered and fully disordered regions of the map.

To label the reference database, DFT reference values for atomic structures generated by AIMD and by iterative training (Fig. 5.1) were computed with CASTEP [349] (version 16.11) using the PBEsol functional and ultrasoft pseudopotentials, yielding energies, forces, and stresses for fitting GAP potentials. A 600 eV cut-off for plane waves and an energy tolerance of 10^{-7} eV per cell were chosen for SCF convergence. A k-point grid of $2 \times 2 \times 2$ was used for large structural models (more than 150 atoms); other computations used an automatically generated k-point grid with a maximum spacing of 0.03 Å^{-1}.

5.2.2 GAP model hyperparameters

All potentials discussed in this chapter were fitted using the Gaussian approximation potential (GAP) framework [164], as implemented in the QUIP / GAP software. A detailed introduction of GAP framework is provided in Section 3.1. Choice of various hyperparameters used in GAP models will be discussed in the following.

Choice of structural representation

GAP is a kernel-based method, using a similarity measure to determine the similarity of any two atomic environments (e.g., the similarity of a new environment to ones "seen" in training). In the simplest case, this is a single kernel function, $k(i, j)$, where $k = 0$ indicates that the i-th and j-th atomic environment are fully dissimilar, and $k = 1$ indicates that they are identical up to the cut-off radius (see

below). However, many GAP models use a combination of multiple descriptors, and the scaling between them as well as their individual settings are important hyperparameters of the model [171]. A detailed discussion of GAP hyperparameters was given in Section 3.1.

For the present work, a combination of squared-exponential two-body (2b) and Smooth Overlap of Atomic Positions (SOAP) many-body descriptors was used. To combine the two, a weighting coefficient, δ, is used to control the contributions from the 2b and SOAP terms (i.e., δ_{2b} and δ_{SOAP}) [171]. To determine suitable values of δ, various potentials were first fitted with different combinations of two δ values, based on the "iter0" datasets, and then the potential-energy curves for atomic dimers were evaluated using these potentials (Fig. 5.2).

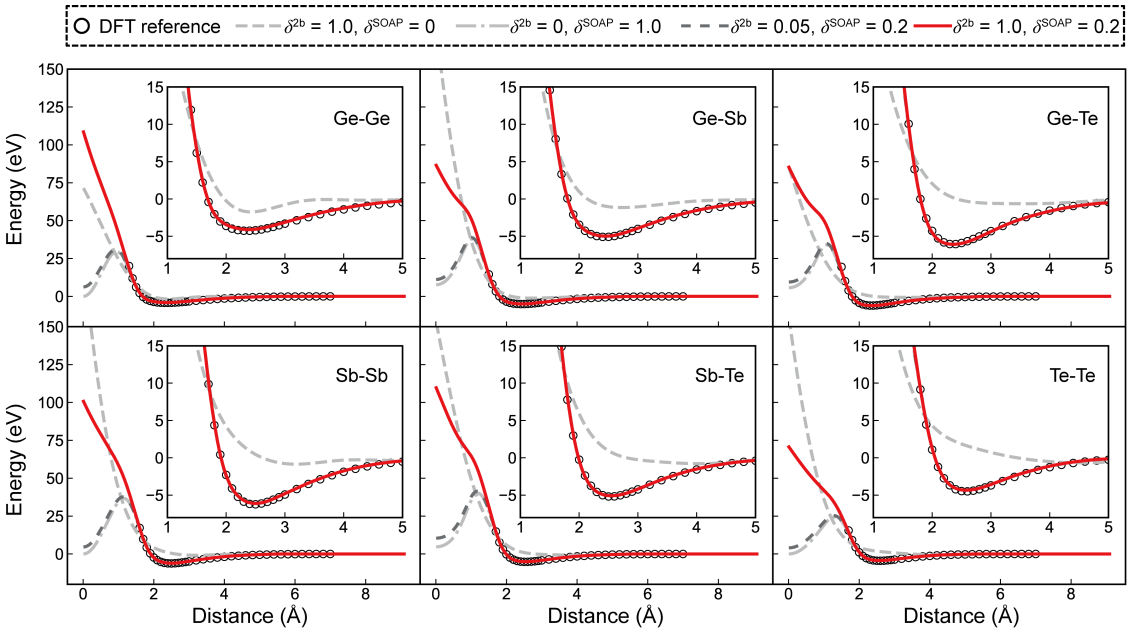


Figure 5.2: Potential-energy scans for atom-pair interactions. Results of isolated-dimer scans for all pairs of atomic species in the Ge–Sb–Te system, viz. Ge–Ge, Ge–Sb, Ge–Te, Sb–Sb, Sb–Te, and Te–Te. Energies relative to the free atoms were computed from various GAP models with different δ values and compared to DFT data.

It is found that models with either a 2b term only (that is, $\delta_{SOAP} = 0$) or a SOAP term only ($\delta_{2b} = 0$) cannot describe the dimer curves for the GST system well. Moreover, even though the combination of a small δ_{2b} (0.05) and a large δ_{SOAP} (0.2) led to a model which can well reproduce the region near the minimum of the

dimer curves, it failed to accurately describe the interactions at small distances (< 1.0 Å). A rapid decay of the dimer curves at small distances was found in this case, which might lead to atomic clustering once atoms are getting too close, especially at high T. (This behaviour is consistent with previous findings for a SOAP-only model for carbon, cf. Ref. 176). Hence, a combination of a large δ_{2b} (1.0) and a small δ_{SOAP} (0.2) was used in this work, and as shown in Fig. 5.2, the resulting model can accurately describe the entire dimer scans in comparison to the DFT reference. (The numerical values given here correspond to those used in the GAP fitting input.)

SOAP hyperparameters

The SOAP structural descriptor is based on an expansion of the atomic neighbour density into a local basis set of radial functions and spherical harmonics, up to a given radial cut-off and with Gaussian smoothing applied to the atomic positions [170]. The corresponding settings for the kernel are important hyperparameters in SOAP-GAP model fitting. The radial cut-off, r_{cut}, was set to 5.5 Å, and the neighbour-density smoothness, σ_{at}, was set to 0.5 Å. For the radial and angular basis functions in the neighbour-density expansion, maximum values of $n_{max} = 12$ and $l_{max} = 4$ were used, respectively.

Notably, the model hyperparameters have been carefully selected, either based on the previously described convergence of the accuracy of GAP models in terms of computational cost (e.g., for n_{max} and l_{max}), as discussed in Ref. 171, or calibrated using the computed potential-energy curves for isolated dimers (e.g., for r_{cut}; Fig. 5.2). In fact, a slightly smaller r_{cut} (e.g., 5.0 Å) will give less smooth and thus inaccurate dimer scans when d is approaching 5.0 Å, whereas the current value (r_{cut} = 5.5 Å) well reproduces the dimer curves. These hyperparameters were also used in the calculation of SOAP structural similarity measures for consistency (e.g., to define the distances visualised in the map in Fig. 5.1b).

Sparsification and regularisation

In the sparsification process (definition of "representative" points for the GAP model fit), different numbers of representative points were used for different parts of the dataset, according to the number of newly added configurations at each iteration. The objective of this scheme is to assign equivalent contributions from each iteration to the final potential, avoiding the overweighting of any specific subset. The numbers of representative points used are summarised in Table 5.1.

For the regularisation (i.e., the "expected error"), which may be thought of as the "expected error" of the input data [171], moderate terms ($\sigma_{\text{energy}} = 10$ meV atom^{-1}, $\sigma_{\text{force}} = 0.2$ eV Å^{-1}, and $\sigma_{\text{stress}} = 0.2$ eV) were first used uniformly for all structures in the reference database (i.e., the main GAP-22 model), to keep the model as general as possible to simulate complex phases of GST, including crystalline, amorphous, and liquid configurations, and intermediate stages of structural phase transitions.

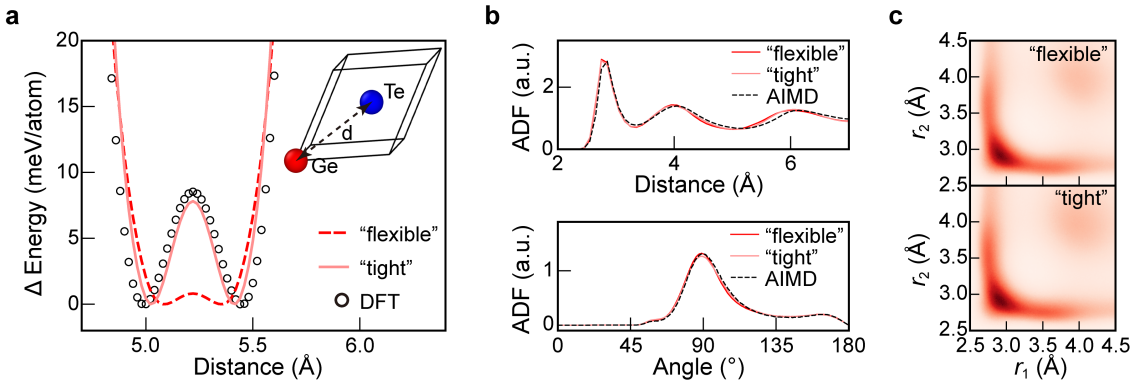


Figure 5.3: Potential-energy scans for atom-pair interactions. Results of isolated-dimer scans for all pairs of atomic species in the Ge–Sb–Te system, viz. Ge–Ge, Ge–Sb, Ge–Te, Sb–Sb, Sb–Te, and Te–Te. Energies relative to the free atoms were computed from various GAP models with different δ values and compared to DFT data.

However, a tight regularisation on the particular part of the reference database might lead to a more accurate description of some subtle properties. For example, to accurately describe the Peierls distortion in crystalline GeTe, the prediction needs to reach very high numerical accuracy, tighter than the regularisation term applied to the main GST-GAP-22 model. A GAP fit using "tighter" energy regularisation for

crystalline structures (2 rather than 10 meV at.$^{-1}$), keeping the reference database and other hyperparameters unchanged, affords an excellent description of the double potential well of trigonal GeTe (Fig. 5.3a) – in other words, high accuracy can be achieved for the Peierls distortion in α-GeTe compared to DFT results. Almost no differences in the ALTBC function, which is used to quantify the Peierls-like distortion as well as short- and medium-range order of amorphous GeTe, were found between the original GST-GAP-22 model and the more tightly regularised fit (Fig. 5.3b–c).

5.3 Validation of ML potentials

The numerical performance of the GST-GAP-22 potential was assessed both by cross-validation on the database itself, and by testing on out-of-sample configurations generated separately via AIMD. The error analysis suggests that the model has indeed "learned" its quantum-mechanical reference data largely to within the regularisation (the imposed uncertainty during the fit).

5.3.1 Numeric validations

To test the numerical accuracy of our GST-GAP-22 model, both "internal" and "external" validation were performed, based on configurations included in and excluded from the reference database, respectively. The computed lattice parameters of various GST crystals using the GAP model were also compared to the DFT reference. The numeric errors of both cross-validation and external validation were shown in Table. 5.2.

Cross-validation (internal validation)

Five-fold cross validation was performed based on the reference database. The training database was randomly split into five equivalent sub-sets ("folds"), and then predicted the cohesive energies and atomic forces of each fold (as the testing dataset), with the remaining folds as the training dataset. The reference database

Table 5.2: Root mean square error (RMSE) results for predicted energies and force components, obtained through both internal (cross-validation) and external (out-of-sample) testing.

		RMSE energies (eV atom^{-1})	RMSE forces (eV Å^{-1})
Cross-validation	Distorted crystals	0.024	0.20
	Disordered AIMD	0.027	0.18
	Standard iterations	0.010	0.18
	Domain-specific iterations	0.007	0.18
Out-of-sample (AIMD structures)	Disordered structures	0.023	0.16
	Crystallisation [114]	0.025	0.15

contains four main types of components, viz. crystalline phases, AIMD data, standard iterations, and domain-specific iterations.

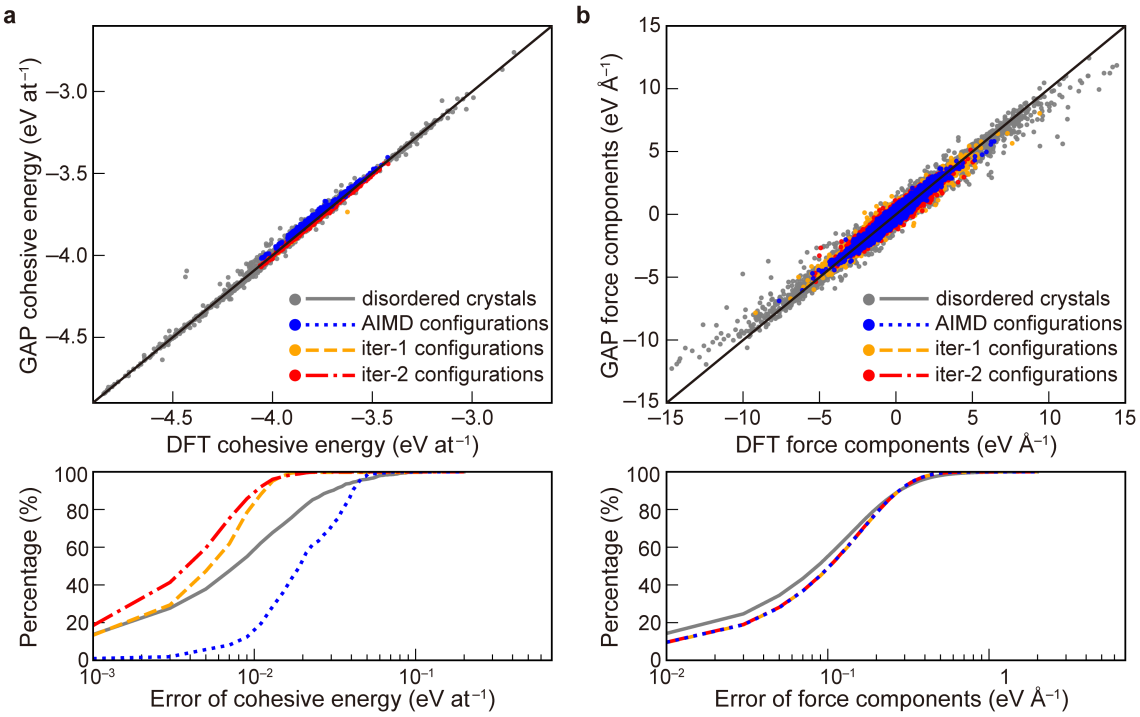


Figure 5.4: Numerical errors of the GST-GAP-22 potential. (a) Scatterplot of cohesive energies, relative to the constituent free atoms, for configurations of different types. The ground-truth DFT value is shown on the horizontal axis, and the prediction of the ML model is shown on the vertical axis. A cumulative error plot, indicating the percentage of data points that have, at maximum, a given error, is given below. (b) Same for errors of the Cartesian force components.

The computed cohesive energies (with respect to the isolated atoms) and force components using GAP are compared to the results obtained from DFT. Most dat-

apoints scatter in a symmetric shape around the diagonal line (Fig. 5.4a–b), which hints at a good numerical agreement between DFT and GAP models overall. Notably, a limited number of crystalline configurations are not as well described by the model (corresponding to the datapoints away from the diagonal line) since they are either highly distorted crystals or rarely visited structures from the Materials Project database [320].

Cumulative error distributions were calculated to assess the performance of the potential for different types of configurations (Fig. 5.5). Almost 60% of the crystalline configurations have an absolute cohesive energy error of less than 0.01 eV atom^{-1}, and for standard and domain-specific iterations this value increases to about 80% and about 90%, respectively. Most energy errors for AIMD-based structures range from 0.01 eV atom^{-1} to 0.1 eV atom^{-1}, slightly larger than for the other three types. The cumulative force error plots (indicating the overall distribution) are rather similar for the four parts of the database, consistent with similar RMSEs (cf. Table 5.2).

External numerical validation

An out-of-sample validation was also carried out, in which the testing configurations were taken from previous AIMD trajectories that were not included in the training datasets. These testing databases consist of two parts. The first one contains 80 disordered GST structures (including Sb_2Te_3, $Ge_1Sb_4Te_7$, $Ge_1Sb_2Te_4$, $Ge_2Sb_2Te_5$, $Ge_3Sb_2Te_6$, $Ge_4Sb_2Te_7$, $Ge_8Sb_2Te_{11}$ and GeTe, with each simulation cell containing more than 300 atoms) obtained from melt-quench AIMD simulations. The second part contains intermediate configurations (1,008 atoms per cell) of the recrystallisation process of $Ge_1Sb_2Te_4$ as described using AIMD simulations in Ref. 114, with data taken from that previous work [114]. The energies and forces were computed using the GAP potential and compared to the AIMD reference (Fig. 5.5). All energy errors for these out-of-sample structures (not included in the training of the model) are lower than 0.5 eV atom^{-1}. Most forces errors are on the order of 0.1 eV Å^{-1},

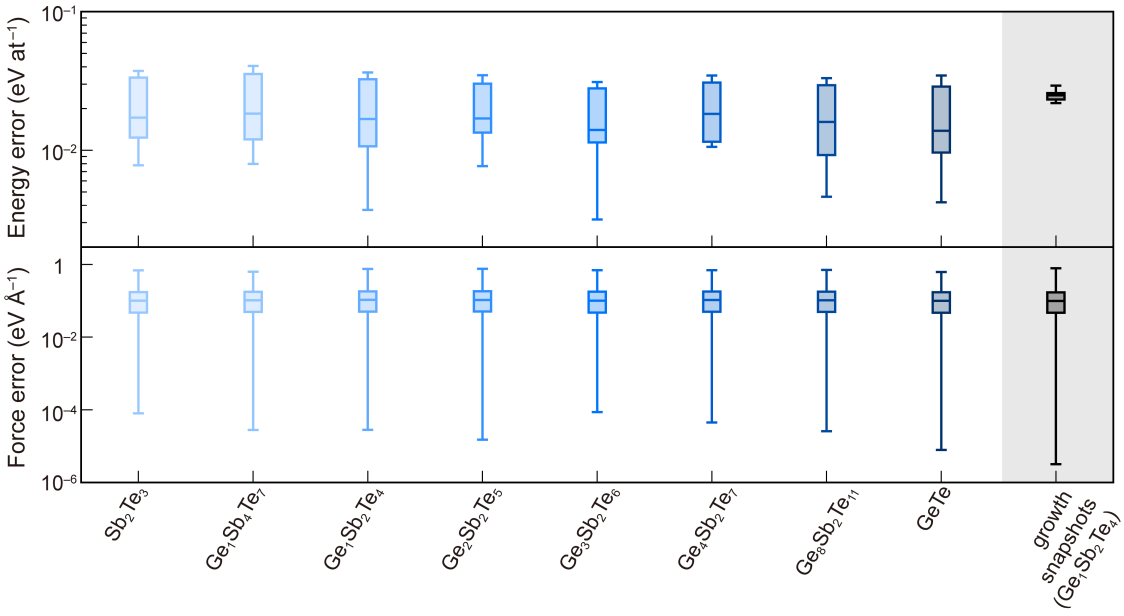


Figure 5.5: Box plots of the errors of cohesive energies and forces, for AIMD configurations of various disordered GST phases (left) and AIMD growth simulations of $Ge_1Sb_2Te_4$ (right). Structural data for the former were generated for the present work, whereas structural data for the latter were taken from Ref. 114. In the box plots, the boxes indicate the range from 25% to 75% with a horizontal line indicating the median, and the whiskers represent the full range of the data.

with some values being as small as 10^{-4} eV Å^{-1}. Given the typical energy difference between the amorphous and crystalline structures of GST (e.g., more than 100 meV atom^{-1}), these errors suggest an accurate description of phase transition processes in GST.

Lattice parameters

Crystalline structures of various GST compounds have been optimised using both DFT and the GAP models. To perform structural optimisation for crystalline GST phases using CASTEP, the plane-wave cut-off energy was set to 600 eV, and the energy tolerance of SCF convergence was 10^{-7} eV per cell. The criterion to break the ionic relaxation loop was a combination of an energy threshold of 10^{-6} eV per cell, a force tolerance of 0.01 eV Å^{-1}, and a stress tolerance of 0.05 GPa. A k-point grid with a spacing of 0.015 Å^{-1} was used to sample reciprocal space. In the structural optimisation using GAP, a force tolerance of 0.01 eV Å^{-1} was used to converge the

optimisation. The obtained lattice parameters show good agreement with the DFT-optimised results (Table 5.3). Maximum relative errors, as compared to DFT, of 1.1% (1.8%) were obtained for PBEsol (PBE) potentials, respectively, indicating the accurate description of the crystalline structures using GAP.

Table 5.3: Computed theoretical lattice parameters of crystalline GST phases, Sb_2Te_3, and GeTe in their respective ground state as obtained by both GAP-MD and DFT computations at 0 K. The difference in lattice parameters stems from the use of different functionals in the DFT computations.

		PBEsol			PBE		
		GAP	DFT	Error (%)	GAP	DFT	Error (%)
Sb_2Te_3	a (Å)	4.252	4.275	0.5	4.345	4.339	0.1
	c (Å)	30.114	29.786	1.1	31.006	31.231	0.7
$Ge_1Sb_4Te_7$	a (Å)	4.253	4.257	0.1	4.331	4.323	0.2
	c (Å)	23.322	23.278	0.2	24.053	24.326	1.1
$Ge_1Sb_2Te_4$	a (Å)	4.254	4.244	0.2	4.321	4.312	0.2
	c (Å)	39.867	40.033	0.4	41.054	41.782	1.7
$Ge_2Sb_2Te_5$	a (Å)	4.245	4.226	0.5	4.310	4.298	0.3
	c (Å)	16.701	16.780	0.5	17.107	17.396	1.7
GeTe	a (Å)	4.173	4.161	0.3	4.242	4.227	0.4
	c (Å)	10.355	10.448	0.9	10.698	10.895	1.8

5.3.2 Structural validations

In addition to numeric accuracy evidenced by the internal and external numeric validations, it is more important for the ML model to also accurately predict the structural properties of GST materials. For crystalline phases, GAP-optimised lattice parameters are nearly identical with DFT data (Table. 5.3). For disordered structures, AIMD simulations were carried out to be the reference. All AIMD simulations shown in this chapter used the "second-generation" Car–Parrinello scheme [325] as implemented in the Quickstep code of CP2K [326]. A combination of triple-zeta plus polarisation Gaussian-type basis set and plane waves with a cut-off energy of 300 Ry was used to expand the Kohn–Sham orbitals and the charge density, respectively. Scalar-relativistic Goedecker pseudopotentials were used [327]. Both the Perdew–Burke–Ernzerhof (PBE) functional [273] and a revised version for solids

and surfaces (PBEsol) [350] were used in the present study. The Brillouin zone was sampled at the Γ-point only, and the timestep was set to 2 fs.

The reference amorphous-phase structural models were generated following a standard melt-quench protocol [111]. Initial configurations were first randomised at 3,000 K (20 ps), then quickly cooled down to the corresponding melting points at a rate of 10^{14} K s^{-1}, and held there for 30 ps. To obtain amorphous structures, the models were then quenched to 300 K at 2.5×10^{13} K s^{-1}, and held there for another 30 ps for data collection.

Given the direct comparison of structural properties between GAP-MD and AIMD, it was found that the GAP models can also faithfully describe the radial and angular distributions of liquid and amorphous GST phases (5.6a-b). Note that the composition $Ge_4Sb_2Te_7$ (formally "4 GeTe + 1 Sb_2Te_3") was not explicitly included during training, yet key structural features of this composition are well described.

All amorphous models contain a non-negligible fraction of homopolar bonds (absent in the respective crystals), e.g., Ge–Ge, typically formed during rapid quenching. The fraction of homopolar bonds is well re-produced across the series of compounds studied (Fig. 5.6c), and so is the energy dependence on the homopolar bond count in a-GeTe, which has been linked to ageing and resistance drift [21]. The onset of a Peierls distortion is a more subtle structural effect in PCMs, which affects the size of the band gap in both crystalline and amorphous GST [21], and is therefore directly relevant to applications in electronic memories. The degree of Peierls distortion can be assessed by comparing the short-to-long bond ratio in close-to linear environments, using a three-body correlation function. Figure 5.6d–e show that the ML potential captures this subtle effect very well: the distortion "finger-prints" for all four compositions studied are similar between GAP-MD (marked in red) and the AIMD reference (marked in blue).

In addition to the compositions "on" the GeTe–Sb_2Te_3 quasi-binary line, the GAP model can also describe different stoichiometric compositions, "off" this quasi-

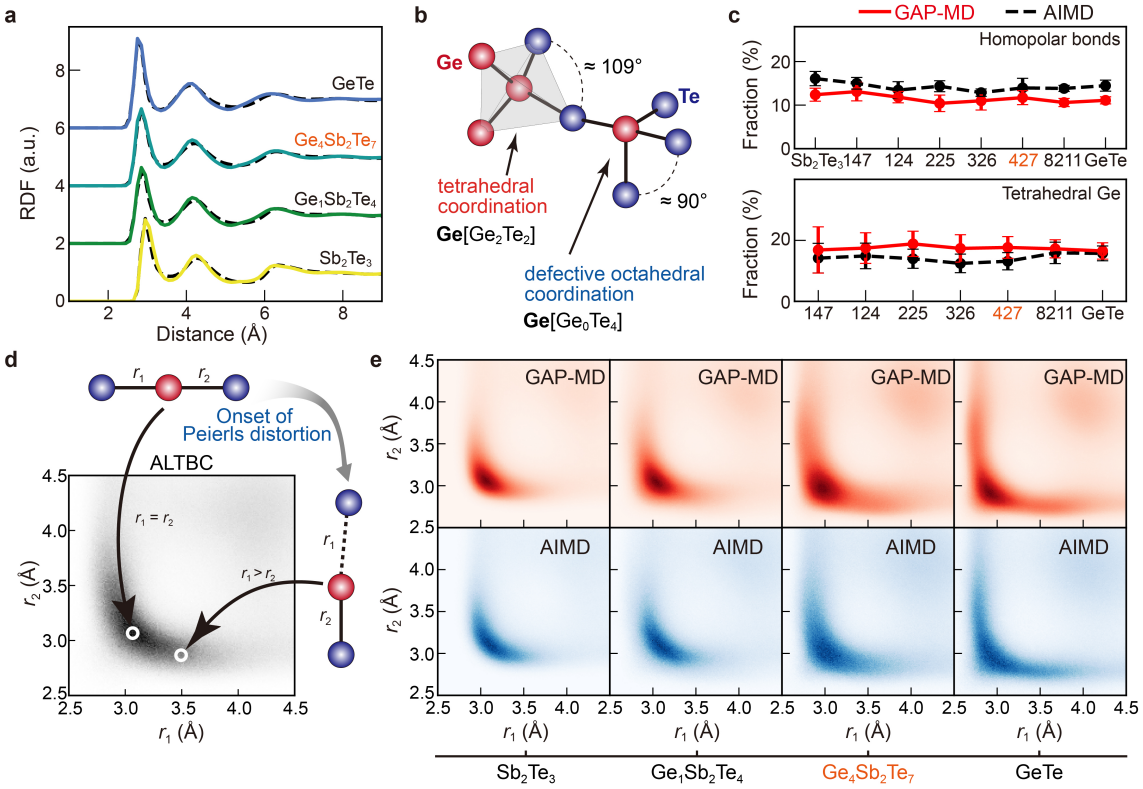


Figure 5.6: Local structure of disordered GST from ML-driven simulations. (a) Radial distribution function (RDF) for four different compositions along the GeTe–Sb_2Te_3 compositional line, from GeTe (cyan) to Sb_2Te_3 (yellow). The $Ge_4Sb_2Te_7$ composition was not directly included in training and is highlighted by red labels. (b) Schematic sketch of octahedral and tetrahedral motifs in amorphous GST. (c) Fraction of homopolar bonds (Ge–Ge, Sb–Sb, Ge–Sb, and Te–Te) and tetrahedrally coordinated Ge atoms in simulated amorphous phases. Results from AIMD (GAP-MD) are shown in black (red), respectively. Error bars indicate standard deviations, and lines between points are guides to the eye. (d) Schematic sketch of the formation of a Peierls-like distortion, the degree of which can be quantified by the normalised angular-limited three-body correlation (ALTBC) function. (e) ALTBC functions for four amorphous GST phases as obtained from GAP-MD (red) and AIMD (blue), respectively.

binary line, to account for the p-type semiconducting behaviour observed in experiments (Fig. 5.7) [76]. Specially, two compositions were considered, including $Ge_2Sb_{1.8}Te_5$ (cation-deficient, p-type) and $Ge_{2.2}Sb_2Te_5$ (cation-rich, n-type) [90]. The structural models of these compositions were generated using AIMD. Structural indicators still yield close similarity between GAP-MD and AIMD (Fig. 5.7): although the above compositions are slightly away from the training domain, their local bonding patterns were "learned" through sufficient sampling over structural and chemical space. All the above analyses demonstrate the compositionally trans-

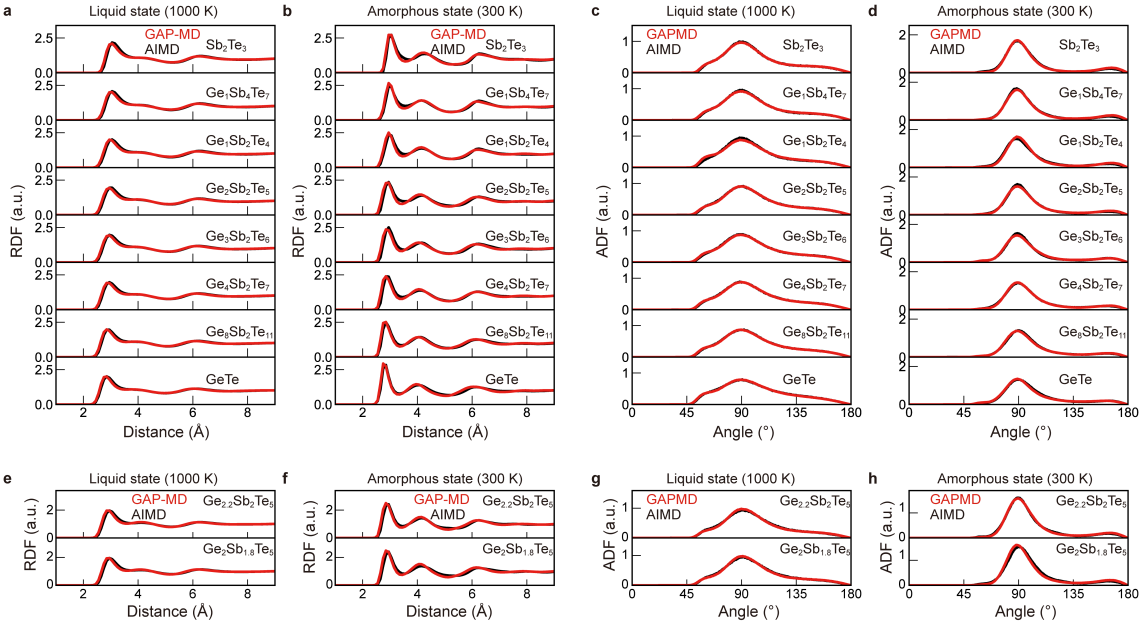


Figure 5.7: Radial and angular distributions. The top row shows calculated radial distribution function (RDF, a–b) and angular distribution function (ADF, c–d) plots of liquid and amorphous GST compounds, obtained using GAP-MD and AIMD at the PBEsol level. Data are shown for 8 compositions on the quasi-binary line, viz. Sb_2Te_3, $Ge_1Sb_4Te_7$, $Ge_1Sb_2Te_4$, $Ge_2Sb_2Te_5$, $Ge_3Sb_2Te_6$, $Ge_8Sb_2Te_{11}$ and GeTe, which were included in the training database, and $Ge_4Sb_2Te_7$ which was not included. The bottom row shows validation results as above, but now for "off-stoichiometric" compositions (not on the quasi-binary line), viz. RDF (e–f) and ADF (g–h) plots for liquid and amorphous $Ge_{2.2}Sb_2Te_5$ and $Ge_2Sb_{1.8}Te_5$, respectively.

ferable and defect-tolerant nature of the current ML model, enabling efficient MD simulations of amorphous GST with local compositional fluctuations—as they are expected to spontaneously occur upon repeated programming operations in memory devices.

More details were given for the medium-range order in amorphous GST and the ageing behaviour in PCMs. The calculated RDFs show an overall agreement of medium-range order of both liquid and amorphous phases of GST (Fig. 5.7). In addition, the distribution of primitive rings of various GST compounds was also computed (Fig. 5.8), showing results similar to AIMD data. Notably, in particular, that the GAP model produced similar fractions of four-membered rings with AIMD, which is the structural building unit of the rock-salt-like phase and is strongly related to the crystallisation kinetics in GST.

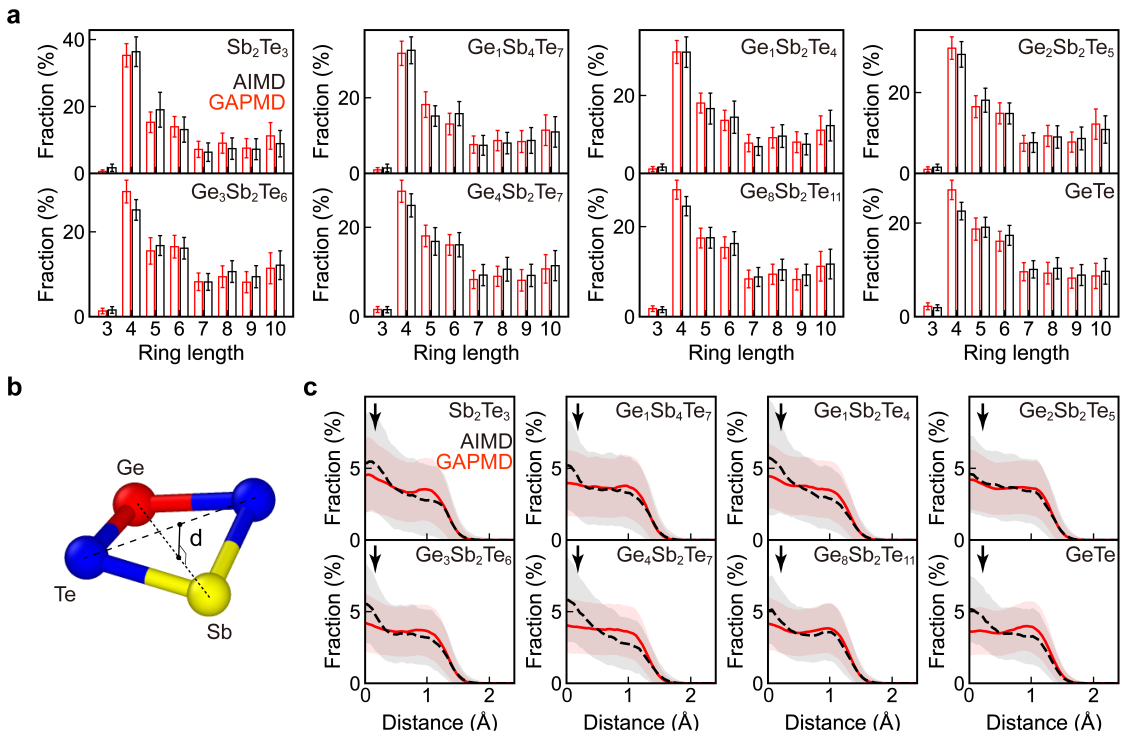


Figure 5.8: Medium-range order. This figure provides further insight into the medium-range order of GST compounds as described by GAP and AIMD models. (a) Distribution of primitive rings. (b) Schematic drawing of how the degree of planarity can be characterised using the inter-diagonal distance, d, of four-membered rings [351]. d values close to 0 indicate overly flat rings. (c) Calculated planarity of various amorphous GST structures from GAP-MD and AIMD. The solid lines present an average value and the shaded areas indicate standard deviations, following Ref. 351.

To go beyond these established measures for MRO, the degree of planarity of all four-membered rings in the AIMD- and GAP-generated structures was further computed, following the same definition as introduced in Ref. 351, and found that the geometric parameters of four-membered rings are same in both GAP and AIMD results. Not too many overly flat rings (which are believed to be the origin of the overly fast crystallisation of GST using other ML potentials [351]) were found, implying consistent crystallisation kinetics between GST-GAP-22-driven MD and AIMD.

In addition, the GAP models also have the potential capacity to study ageing behaviour in GST at the atomic scale. Previous work, based on AIMD simulations shown in Ref. 21, showed that an elemental substitution protocol, i.e., replacing

Si (or Sn) atoms with Ge in melt-quenched SiTe (or SnTe), can generate diverse amorphous networks with more (fewer) tetrahedral Ge local environments, respectively, compared to the conventional melt-quenched amorphous GeTe models (Fig. 5.9a). A clear ageing path, i.e., an increase of the calculated optical band gaps of these amorphous samples, was observed (Fig. 5.9b), indicating that a structural relaxation process from tetrahedral Ge-rich environments to tetrahedral Ge-poor environments leads to more energetically favourable configurations (Fig. 5.9c).

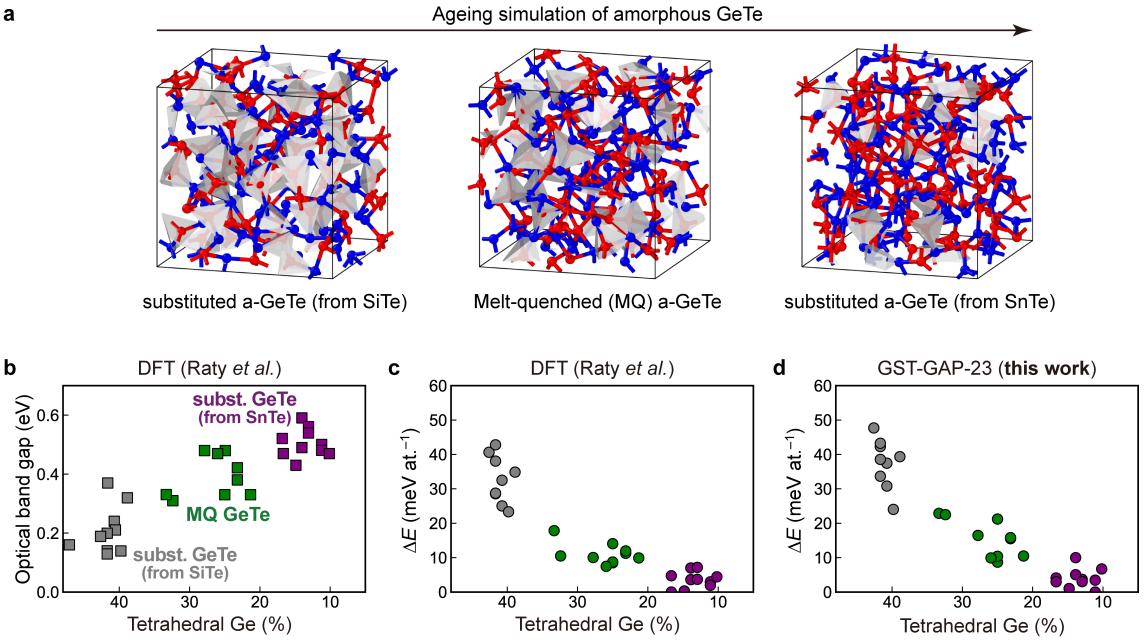


Figure 5.9: Energetics of ageing in a-GeTe by DFT and GST-GAP-22. (a) Different amorphous structures of GeTe. Elemental substitution (i.e., replacing Si or Sn with Ge in melt-quenched SiTe or SnTe) can be used to generated amorphous GeTe with more (if generated from SiTe) or fewer (from SnTe) tetrahedral Ge motifs. (b) Calculated optical band gap and (c) energy difference of these amorphous GeTe structures using DFT. Data are collected and replotted from Ref. 21. (d) As panel c, but recomputed using GST-GAP-22, providing evidence that this model can also capture the subtle energy difference of the same amorphous structures as compared to the DFT results in panel c.

To test the behaviour of the GAP model for the aforementioned ageing behaviour, the GST-GAP-22 potential was used to re-compute the energies of these amorphous samples, and a same energetic ranking was obtained (Fig. 5.9d). This result suggests that the GST-GAP-22 models can capture the very small energy differences (at the level of tens of meV at.$^{-1}$) between various amorphous configurations of GeTe, and

that they are able to characterise the energetics of the structural evolution during the ageing process. It is expected that in the future, advanced structural-search protocols, e.g., Monte Carlo runs based on ML local energies [331], can be applied together with our GAP models to study the aging trend of various GST compounds in larger model systems (including up to device scales), which goes beyond the computational capacity of DFT methods.

5.4 Modelling conditions to represent devices

Simulating the realistic programming conditions in GST-based devices is critical for real device simulations and serves as a big challenge. Figure 5.10 shows three typical operations that need to be represented in simulations. Amorphous PCM films can be crystallised either by isothermal annealing or by Joule heating (programming pulse). In thin-film experiments, amorphous GST alloys crystallise into disordered rock-salt-like phases upon substrate heating at $\approx$ 150 ℃, accompanied by a density increase (volume shrinkage; Fig. 5.10). Further annealing (250-400 ℃) induces a gradual structural change via continuous vacancy ordering [11, 76]. Such annealing is the initial step both for electronic and photonic PCM devices. The subsequent operation is to melt a fraction of the active volume near the heater by electrical or laser pulses, and the following rapid cooling freezes the material to form the amorphous phase. In photonic devices that feature a more open structure, the active volume expands and shrinks frequently; in electronic devices, the PCM is fully encapsulated, and the active volume remains nearly unchanged upon programming.

Describing these different scenarios leads to different use cases for ML-driven MD (MLMD) conditions: namely, simulations in the canonical ensemble (NVT, constant number of particles, volume, and temperature), the isothermal–isobaric ensemble (NPT, constant number of particles, pressure, and temperature), and the microcanonical ensemble (NVE, constant number of particles, volume, and energy). The timestep was set to 2 fs in all MLMD simulations (the same as in the AIMD runs). A Langevin thermostat (using the default damping value in LAMMPS) was

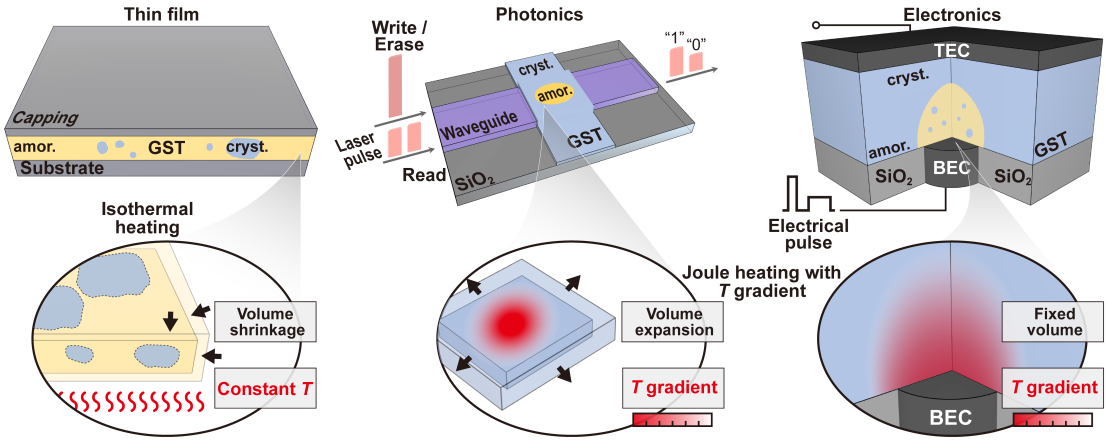


Figure 5.10: Schematic sketches showing, from left to right: thin-film laboratory samples, photonic devices [342], and electronic devices [47]. Insets indicate different programming conditions in GST, the MD simulation of which requires different ensembles (NVT, NPT, and NVE).

used to control temperature in the NVT simulations, and a Nosé–Hoover thermostat [352, 353] controlling temperature and a barostat [354] controlling external pressure were used in the NPT simulations.

Given the statistical distribution of vacancies in rock-salt-like GST and the highly disordered nature of liquid and amorphous GST, MD simulations of a few hundred atoms are naturally affected by strong finite-size effects. In parallel simulations of heating small-scale rock-salt-like $Ge_1Sb_2Te_4$ models (Fig. 5.11), the SOAP-based crystallinity, $\overline{k}$ [114], falls at different onset temperatures for five independent runs with different initial atomic configurations (NVT ensemble), and large fluctuations in mass density are seen when the volume is allowed to change (NPT). In contrast to AIMD, GAP-MD enables efficient simulations for much larger model sizes, improving statistical sampling and thereby the description of the systems. In particular, the $\overline{k}$ and mass-density profiles nearly overlap for $12 \times 12 \times 12$ supercells (12,096 atoms each) over independent NVT and NPT runs. The same holds for NVE simulations, yielding more consistent liquid dynamics in larger supercells (Fig. 5.11). These results suggest marginal finite-size effects in larger supercells, in line with previous simulations of glassy $Ge_2Sb_2Te_5$ [154]. To better illustrate the usefulness of the

GAP models in the simulations of device environments and device operations, more demonstrations will be given in the following sections.

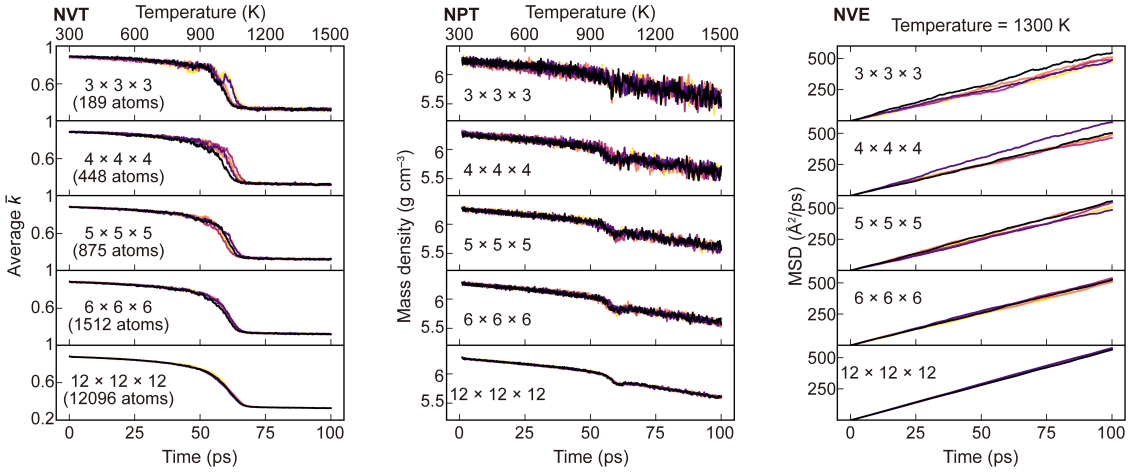


Figure 5.11: Evolution of relevant quantities during GAP-MD simulations of $Ge_1Sb_2Te_4$ with different system sizes, expressed as $n \times n \times n$ expansions of the cubic rock-salt-type unit cell. From left to right, different results were shown, including the per-cell averaged crystallinity measure, $\overline{k}$, during NVT heating, the mass density during NPT heating, and the mean square displacement (MSD) during NVE MD at 1,300 K (with initial configurations taken from the NVT runs). For each model size, five independent simulations were performed, and the results are plotted as semi-transparent lines.

5.4.1 Single and cumulative switching

The first example of using GAP models in simulating device-related scenarios is the crystallisation (SET) process. Figure 5.12a shows a state-of-the-art AIMD crystallisation simulation of $Ge_1Sb_2Te_4$ over 600 ps at 600 K, with colour-coding for the per-atom similarity $\overline{k}$ (data from Ref. 114). The model contains 1,008 atoms ($3.0 \times 2.9 \times 4.2$ nm^3), with two fixed crystal-like layers representing the crystal–amorphous boundary encountered in devices (Fig. 5.12a). These AIMD simulations have been discussed in detail in Section 3.5. Note that this is almost the largest size for AIMD simulations. The new GAP-MD simulation successfully reproduces the crystallisation process (Fig. 5.12b), giving a very similar crystallisation process, evidenced by the colour-coding using $\overline{k}$. In addition, the resulting structure similarly contains a random distribution of vacancies and a number of anti-site defects. Note that this particular type of GAP-MD run can be completed within days on a high-

performance computing system, whereas the corresponding AIMD simulation [114] required half a year of real time and shows much less favourable scaling behaviour.

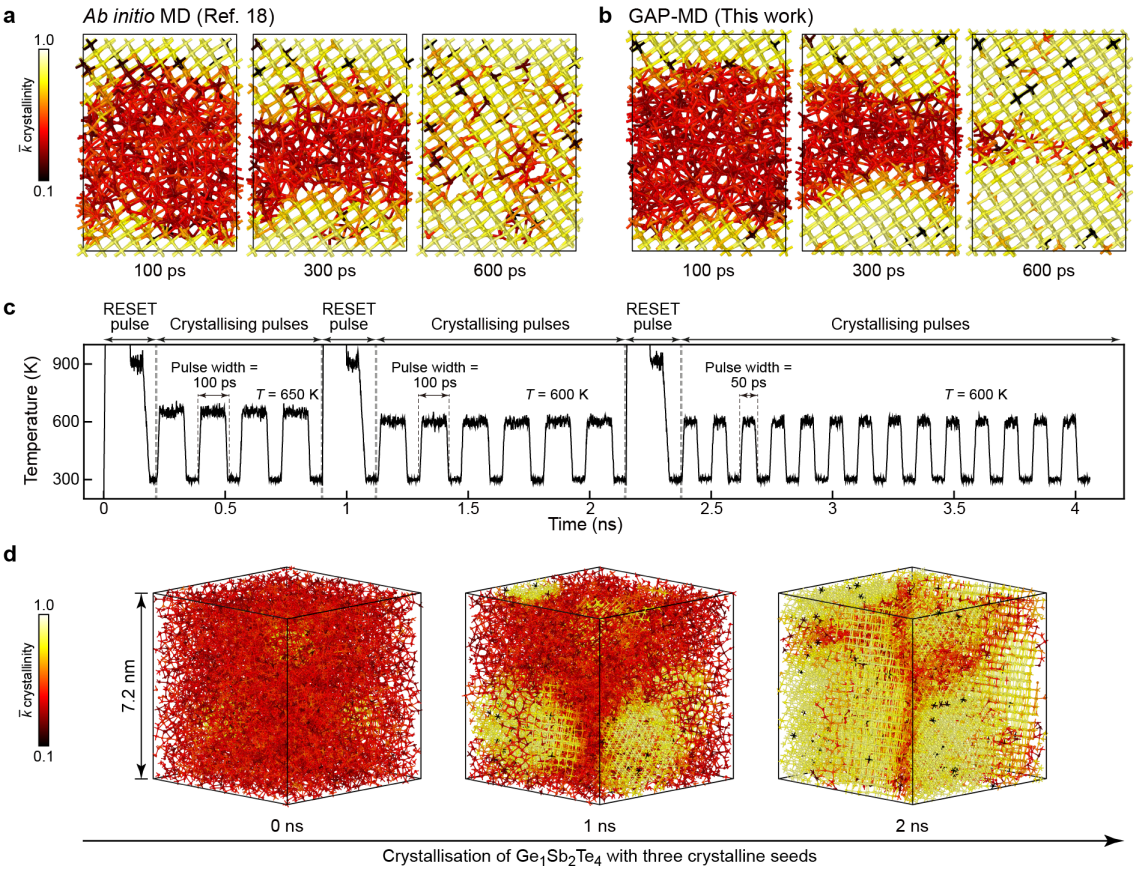


Figure 5.12: Thermal cycles and cumulative SET operation. (a) Snapshots from a 1,008-atom constant-temperature AIMD crystallisation trajectory for $Ge_1Sb_2Te_4$ in the <111> direction, taken from Ref. 114 and visualised in the same style. Colour-coding indicates the SOAP-based $\overline{k}$ crystallinity [114], illustrating the gradual transition from amorphous (red) to crystalline (yellow), and the occurrence of anti-site defects (black). (b) The same crystallisation simulation now carried out using GAP-MD. (c) GAP-driven simulations of cumulative SET processes. After each RESET (amorphising) pulse above the melting point of $Ge_1Sb_2Te_4$, a series of small SET (crystallising) pulses were applied, with the same simulation setup as shown in panel b. In this proof-of-concept, three different kinds of pulses with different amplitudes and durations were considered: 650 K pulses of 100 ps duration, 600 K pulses of 100 ps duration, and 600 K pulses of 50 ps duration, respectively. The lower panel characterises the resulting cumulative SET processes by the averaged $\overline{k}$ crystallinity [114] as a function of the number of applied pulses. Each data point corresponds to a configuration after an individual pulse, and the lines are only guides to the eye; the starting point of the plot (corresponding to 0 ns) refers to the fully recrystallised configuration from the simulation shown in panel b. (d) Crystallisation of $Ge_1Sb_2Te_4$ with multiple nuclei, leading to formation of a polycrystalline structure over 2 ns.

In addition to SET and RESET, two further operations—cumulative SET and iterative RESET—are used for neuro-inspired computing [340]. In these cases, the

programming pulse can be divided into a train of pulses with varied amplitude and duration [115], and the change in electrical resistance [31] or optical transmission [68] as a function of the number of pulses can be used to mimic synaptic learning rules. Using the GAP model, the cumulative SET simulations were carried out (Fig. 5.12c). Differences in the initial configuration, pulse width, and maximum temperature all affect the progressive structural changes, and the evolution of structural order determines the electronic and optical properties of the intermediate states. These simulations represent the situation inside one grain of rock-salt-like GST (with the grain size varying from several to several tens of nanometres) [13, 355]. Taking into account the randomness of nucleation of GST, cumulative SET operations can emulate integrate-and-fire dynamics for the construction of stochastic phase-change neurons [356]. The simulations shown in Fig. 5.12a and b are at the system-size limit of DFT: therefore, they describe crystal growth in a necessarily simplified way, along one selected direction only, viz. [111].

However, the full crystallisation process of GST mainly consists of two different parts, viz. nucleation and growth [1]. Because nucleation is dominant in GST, it takes several to tens of nanoseconds to nucleate, due to the largely stochastic nature of the incubation process. To demonstrate this stochasticity, ten independent nucleation simulations of $Ge_1Sb_2Te_4$ were performed at 550 K using the GAP potential (Fig. 5.13). Only 7 out of 10 structural models were crystallised within 10 ns, evidencing the large stochasticity (Fig. 5.13). As soon as this incubation process is finished, the subsequent crystal growth can be completed within a few nanoseconds or even hundreds of picoseconds [108].

Based on the capability to accurately simulate the nucleation process in GST, the GAP potentials make it possible to describe a full crystallisation with growth processes along different orientations: in Fig. 5.12d, a larger-scale crystallisation simulation was exemplified to obtain a polycrystalline state, which is typically formed in nanoscale devices. Three crystalline seeds (of about 100 atoms each) were em-

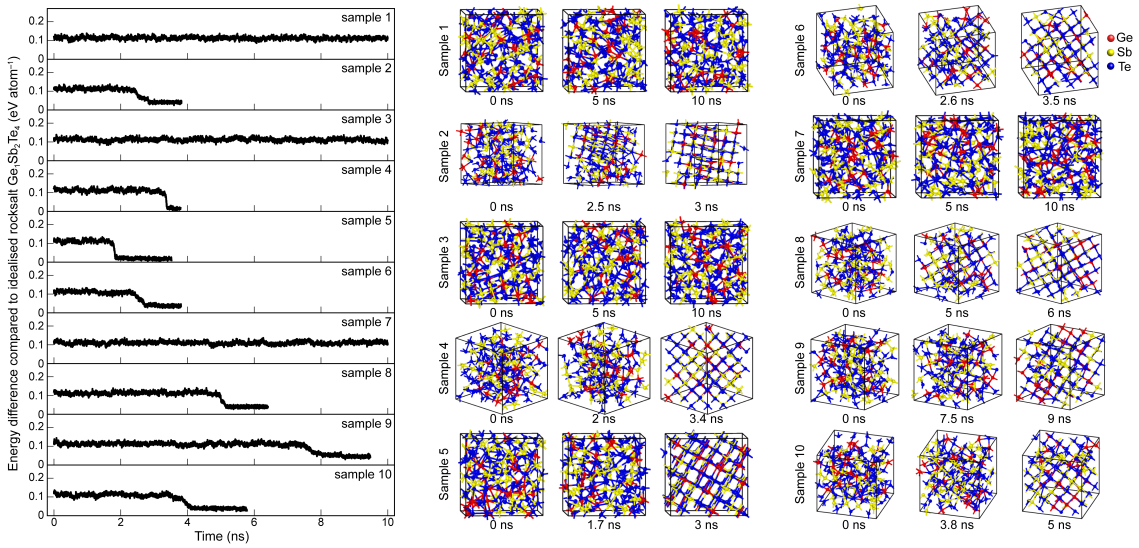


Figure 5.13: Simulated nucleation of $Ge_1Sb_2Te_4$. The plots on the left show the computed energy profiles of ten independent nucleation simulations. The starting configuration is a crystalline rock-salt-like $Ge_1Sb_2Te_4$ ($3 \times 3 \times 3$ supercell, containing 189 atoms). The structural models were first melted and quenched to remove any crystalline order. They then started to crystallise for at most 10 ns at 550 K. Snapshots of nucleation processes (or absence thereof) for ten samples are shown on the right.

bedded in this model, bypassing the stochastic nucleation process (Fig. 5.13). The extracted growth speed averaged over three grains is ≈ 0.7 m s^{-1} at 600 K, consistent with ultrafast differential scanning calorimetry experiments [281]. The large-scale GAP-MD simulations point towards a way to conceptually connect between the previously shown small-scale models and the following real device-size simulations.

5.4.2 Non-isothermal heating

Different types of scenarios in real devices require different modelling strategies, as illustrated in the previous Section 5.4 and Fig. 5.10. The SET process in encapsulated electronic cells is commonly described in isothermal simulations (NVT ensemble), as exemplified in Fig. 5.12. In contrast, simulating the iterative RESET process requires the use of large-scale constant-energy simulations (NVE ensemble). The widely used mushroom-type electronic device was used to model such an operation (Fig. 5.10). To this end, a $32 \times 5 \times 20$ supercell expansion of a rock-salt-like $Ge_1Sb_2Te_4$ model was constructed (which contained 12,100 atoms), and applied en-

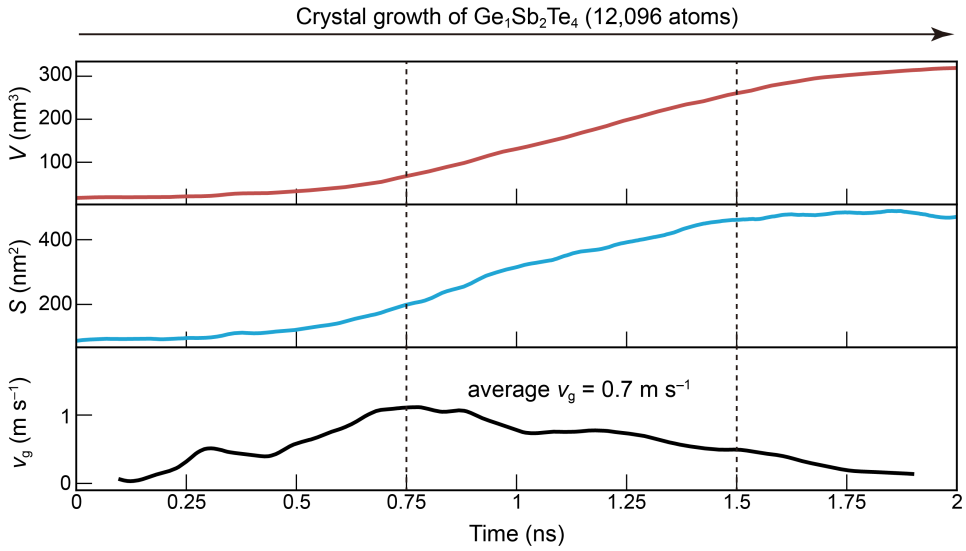


Figure 5.14: Crystal growth in a large model containing multiple nuclei (as a supplement to Fig. 5.12d). The volume of the crystalline area, V, the area of the amorphous–crystalline interface, S, and the crystal growth velocity v_g as a function of crystallisation time for the 12,096-atom model, calculated using the SOAP crystallinity $\overline{k}$. see Ref. 114. A correction for "noise" atoms was performed before calculating these quantities, as discussed in detail in Ref. 114. An averaged growth velocity ($\approx$ 0.7 m s^{-1}) at 600 K was calculated in the time window in which steady growth was observed (dashed vertical lines).

ergy pulses to an area described by a half-cylinder shape at the bottom of the xz plane (Fig. 3b). This setup represents the middle slab of mushroom PCM cells, with a cross-section of $\approx 20 \times 12$ nm^2. The top four atomic layers were fixed to prevent unwanted thermal transport to the top area through the periodic boundary conditions: in a real device, there would be a thermal barrier in contact with the GST material. Note that in state-of-the-art mushroom-type devices [355], the diameter of the heater can be scaled down to as small as $\approx$ 3 nm. Therefore, the expected effective heating area is directly mirrored by our simulation setup that is characterised in Fig. 5.15.

After a short NVT simulation at 300 K, the ensemble was switched to NVE and additional energies of 55 eV ps^{-1} were imposed on the atoms as kinetic energy inside the focal area over 30 ps. The NVE simulation was continued for another 70 ps. In total, this 30 ps energy pulse added 1,650 eV (0.264 fJ) of thermal energy to the model, resulting in a disordered area with a base length of $\approx$ 7 nm (Fig.

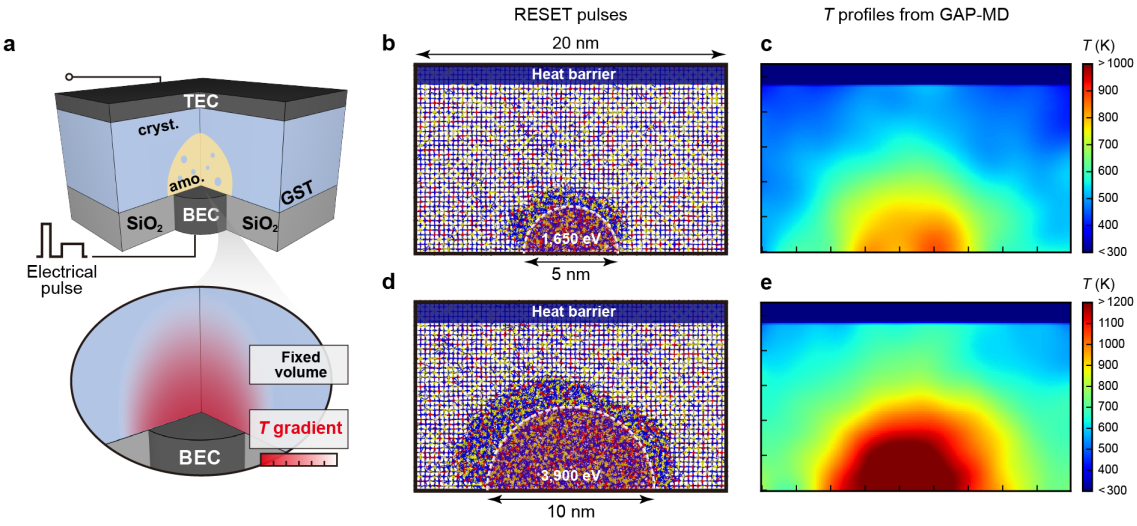


Figure 5.15: RESET operation in a mushroom-type device setup. (a) Schematic sketch of the device layout. (b) Snapshot from a non-isothermal (NVE) simulation in which an additional energy of 1,650 eV in total was imposed on the atoms as kinetic energy inside the focal area over the first 30 ps. A snapshot of the NVE simulation at 100 ps is shown, with the initial focal area indicated by dashed lines. (c) The corresponding two-dimensional temperature profile. The data were smoothed using a Gaussian filter with a broadness of 1.0 nm. (d–e) As panels b and c, respectively, but now for an initial kinetic energy of 3,900 eV.

5.15b). Then, the local temperature profile was computed based on the velocity of atoms collected on a two-dimensional grid. An averaged temperature was computed for each grid, based on the per-atom velocities in this grid. To get a more smooth temperature gradient with less statistical noise, a Gaussian filter was used with a smoothness value to smear out the numerical fluctuation. In this case, the smoothness value was also calibrated (Fig. 5.16).

Figure 5.15c shows a clear radial temperature gradient from the heating centre towards the outer shell. In parallel, another NVE calculation was performed with an energy pulse of 3,900 eV (0.625 fJ) imposed on a larger focal area over 30 ps (Fig. 5.15d). The stronger pulse generated a larger disordered area with a base length of $\approx$ 14 nm, giving rise to a more extended gradient profile with higher average temperature (Fig. 5.15e). This ML-driven MD simulation protocol closely resembles the iterative RESET operation in devices, providing real-time atomic-scale information beyond what has been revealed in experimental observations as well as

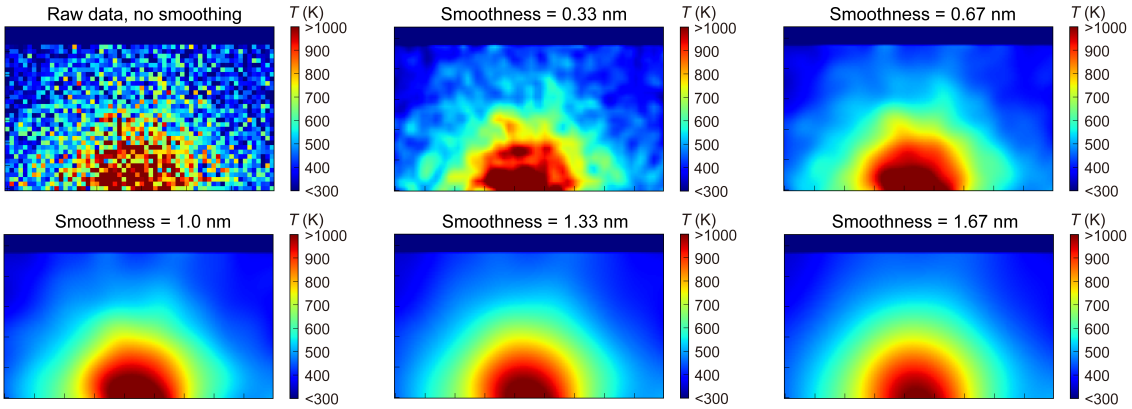


Figure 5.16: The plots provide more detail regarding the calculated temperature profile of the mushroom model during the non-isothermal heating process (Fig. 5.15). (a) The cross-section (perpendicular to the plane shown) of the model was first divided into a 60 × 36 grid (21,600 pixels in total). The per-atom temperature was calculated as an average over all atoms in each grid pixel, leading to the raw data shown in the top left panel. Gaussian filters with different smoothness values can be used to smear out the statistical fluctuation of calculated temperatures. After smoothing, the resolution was further improved by linear interpolation using a finer grid.

finite-element method (FEM) simulations [25]. The evolution of the temperature gradient can be traced not only with sub-nanoscale spatial resolution, but also with an ultrafine time resolution at the femtosecond level (Fig. 5.17).

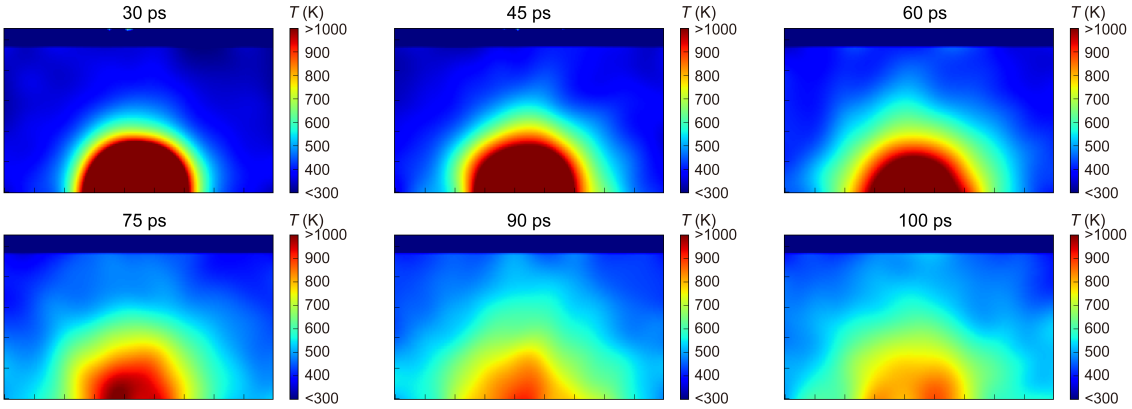


Figure 5.17: Snapshots of the evolution of the temperature gradient during non-isothermal heating. After a 30-ps heating pulse, another 70-ps NVE simulation was performed. A Gaussian filter with a smoothness of 1.0 nm was applied to the data.

5.4.3 Unlocking device-scale simulations

Lastly, a capability demonstration was presented for a structural model of real device size, corresponding to a 3D cross-point product as currently used in electronic memories [8]. Figure 5.18a shows the schematic layout of the crossbar array, which consists of a PCM memory layer (composition close to $Ge_1Sb_2Te_4$), an Ovonic threshold switching (OTS) selector layer, and several buffer layers between PCM, OTS, and electrodes. The rest of the volume is filled with dielectric materials [8]. When programming, the electric current flows through the OTS layer after reaching a threshold voltage, and induces Joule heating to trigger the phase transition of the PCM. The size of the PCM region was reported as $\approx 20 \times 20 \times 40$ nm^3 [8]. The initial state is taken to be the layered trigonal phase due to unavoidable thermal annealing during the backend-of-the-line (BEOL) stage for PCM integration [25]. A $Ge_1Sb_2Te_4$ supercell of the same size was constructed, containing more than half a million atoms in a layered configuration (30 septuple-layer blocks, with each block arranged in the stacking sequence Te–Sb–Te–Ge–Te–Sb–Te), and melted the model using a short (10 ps) but very-high-energy pulse with a high energy gradient. Such a short time window is experimentally accessible in principle: melting of GST can be achieved in 5–10 ps, as evidenced by femtosecond laser experiments [357].

Figure 5.18b illustrates the progressive disordering during melting. Detailed analysis of the changes in atomic structure is possible for such atomistic simulation: layered structures remain early on (clearly seen in the plot at 2 ps). Upon further heating, chemical disordering proceeded vertically from the bottom to the top, resulting in a gradient of ordering (Fig. 5.18b). At 6 ps, the bottom slab still shows some remaining chemical order, but it appears entirely disordered after 10 ps. The top region, especially the structural building blocks close to the heat barrier (near $z \approx 40$ nm), only partially melted after 10 ps.

In addition, the calculated vertical profiles of elemental positions (Fig. 5.19) show that the melting of Ge layers starts slightly earlier than that of the Sb ones,

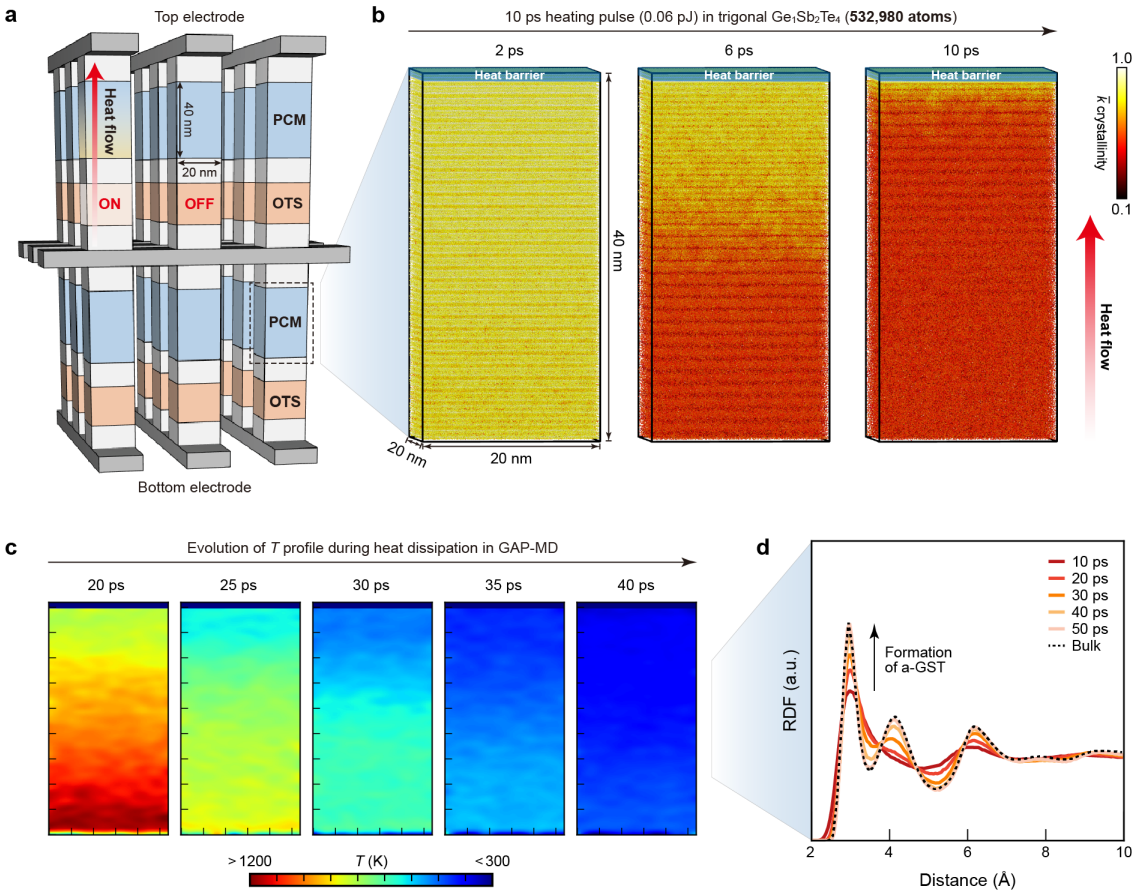


Figure 5.18: Melting and quenching in a confined device setup. (a) Schematic layout of commercial 3D Xpoint products, consisting of electrodes, PCM, OTS, and buffer layers in between, drawn based on information in Ref. 8. (b) Melting process of a $Ge_1Sb_2Te_4$ structural model that matches the dimensions of the PCM cell in 3D Xpoint memory (20 × 20 × 40 nm^3, containing 532,980 atoms). The melting process is triggered by a vertical heating pulse with additional energy (containing 60 fJ in total) from the bottom to the top after the cell is programmed. The colour coding is now indicating the gradual loss of crystalline order (*yellow*) upon formation of a disordered phase (*red*) in a much larger simulation cell. (c) Heat dissipation during an extended MD simulation. (d) Evolution of the a-GST structure from the melted state, characterised using radial distribution function (RDF) plots for the central area of the model (central 20 nm). This figure illustrates how atomic-structure analyses that are common for small cells (cf. Fig. 5.12; "Bulk") can now be performed for full device-scale simulations.

and the Te atoms near the van-der-Waals-like gaps are more mobile than those in the centre of the structural building blocks. At 6 ps, the bottom slab still shows some remaining chemical order, but it is entirely disordered after 10 ps. The top region, especially the structural building blocks close to the heat barrier (near z ≈ 40 nm), only partially melted after 10 ps.

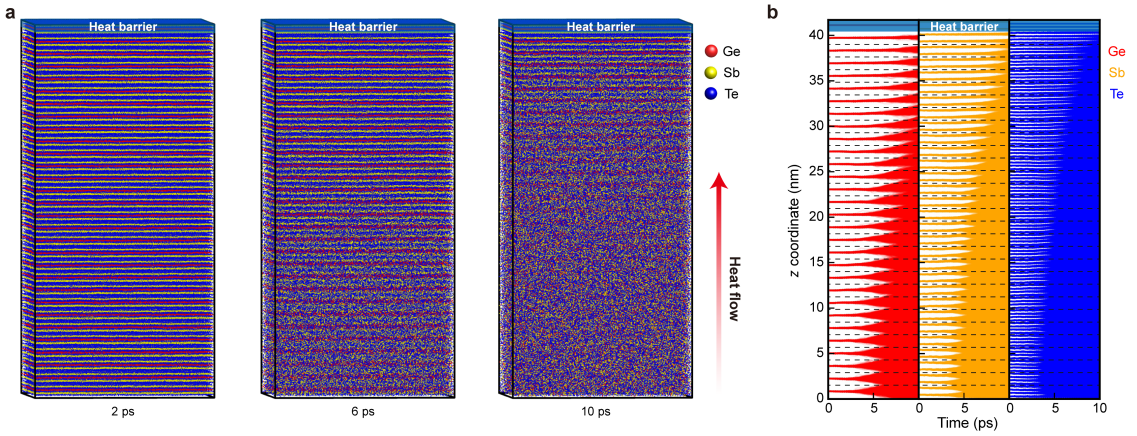


Figure 5.19: Melting in a confined device setup. (a) Melting process of a $Ge_1Sb_2Te_4$ structural model but now with colour-coding according to atomic species, emphasising the gradual melting of the layered GST structure. (b) Time-resolved evolution of the z-axis coordinates of the Ge, Sb, and Te atoms in the system.

Note that in electronic devices, the applied heat dissipates mostly through electrodes (of very high thermal conductivity), quenching the molten state to form the amorphous phase. Here, the added kinetic energy was gradually taken out from each atom over another 40 ps (Fig. 5.18c), leading to a gradual formation of the amorphous phase (Fig. 5.18d). Based on ML-driven MD, the evolution of temperature gradients with sub-nanometre spatial resolution can be obtained (Fig. 5.18c), revealing finer atomic details of the heat-dissipation process as compared to coarse-grained FEM simulations.

The RESET energy is calculated as 60 fJ, which could be used as the theoretical limit of minimum power consumption for the initial operation per bit (the most power-consuming step in devices). This is still an approximation because the entire amount of energy was directly as-signed to the atoms with no energy dissipation to the surrounding layers. Whilst the present capability demonstration focuses on the initial ground state, similar calculations and analyses can now be carried out for more complex structures, including realistic models of the rock-salt-like phase with grain boundaries or mixed crystalline / amorphous structures. Moreover, the simulated heating pulse can be adjusted to more realistic and complex programming conditions, considering various pulse amplitudes, durations and shapes, which shall

help to guide the optimisation of programming schemes for better device performance.

5.5 Towards a complete reference database

As shown in previous sections, the diverse GST-GAP-22 reference database has enabled accurate description of various phases of GST alloys as well as realistic programming scenarios in real-world devices. This reference database has covered a wide range structural configurations, but can still be further extended according to the specific requirements. In this section, the transferability of the GST-GAP-22 database is discussed.

5.5.1 Relabelling at a different DFT level

The aforementioned GST-GAP-22 potential was fitted to reference data obtained with the PBEsol exchange–correlation functional, however, it is shown here that a version based on a different functional can also be generated by re-computing the DFT data labels.

To demonstrate this, a new version of the model that was "re-fitted" at the PBE level was provided. Instead of repeating the entire iterative training process, single-point DFT computations were carried out using the PBE functional for the atomic structures included in the reference database. The obtained DFT energies, forces, and stresses were then used to fit the PBE version of the model. This version reproduces AIMD data at the PBE level with good accuracy, as shown in Fig. 5.20. Note that there exists a difference in the computed lattice parameter (and thereby in the mass density), depending on which functional is used: PBEsol leads to a larger density than PBE. Nevertheless, in both cases, our GAP computations reproduce the DFT ones at the corresponding level (Table 5.3).

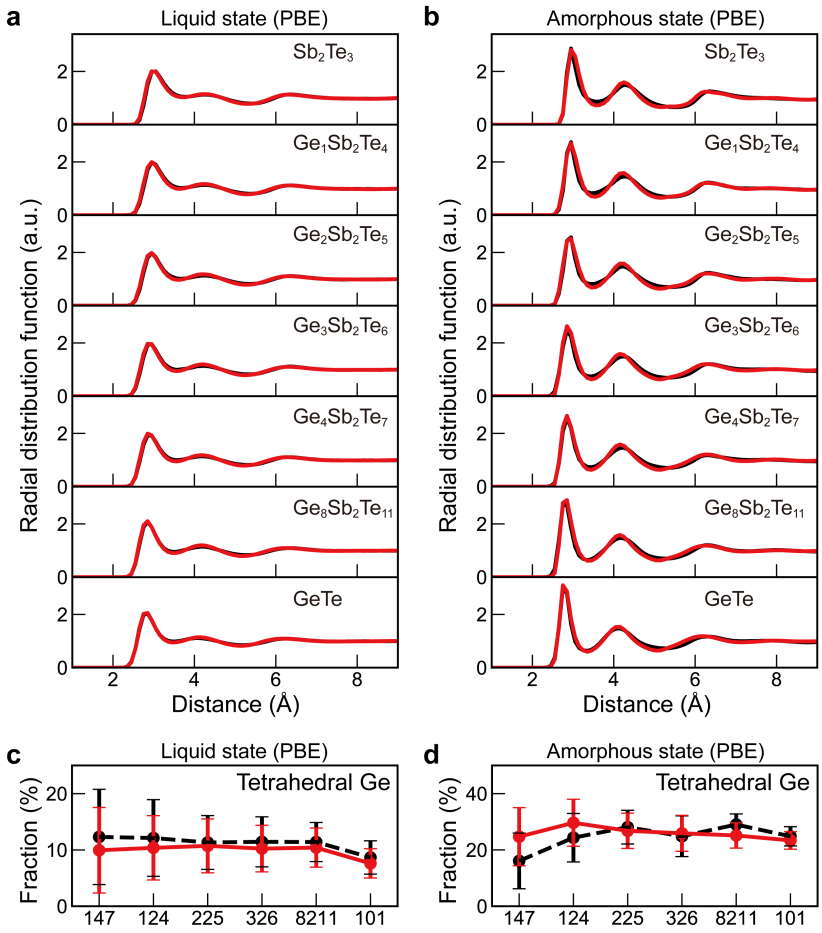


Figure 5.20: Refitting the potential at a different DFT level. Structural validation for a version of the GAP model fitted using PBE reference data. (a–b) Computed radial distribution function (RDF) plots for liquid and amorphous GST structures obtained from GAP-MD in comparison to AIMD, both based on the PBE functional. (c–d) The fraction of tetrahedrally coordinated Ge atoms in liquid and amorphous GST, calculated from GAP-MD (PBE) in comparison to AIMD (PBE) data.

5.5.2 Further extensions - electromigration in GST

The GST-GAP-22 database is designed to be widely applicable as it stands, although it can be further extended if required. An example is discussed here in the context of electromigration, which leads to gradients in the average distribution of cationic- and anionic-like atoms in the system. There are two possible approaches to this computational task, which was sketched in Fig. 5.21a: one may either use the existing database (grey) out of the box, or one may extend it with task-specific structural snapshots that mirror the new application scenario (green).

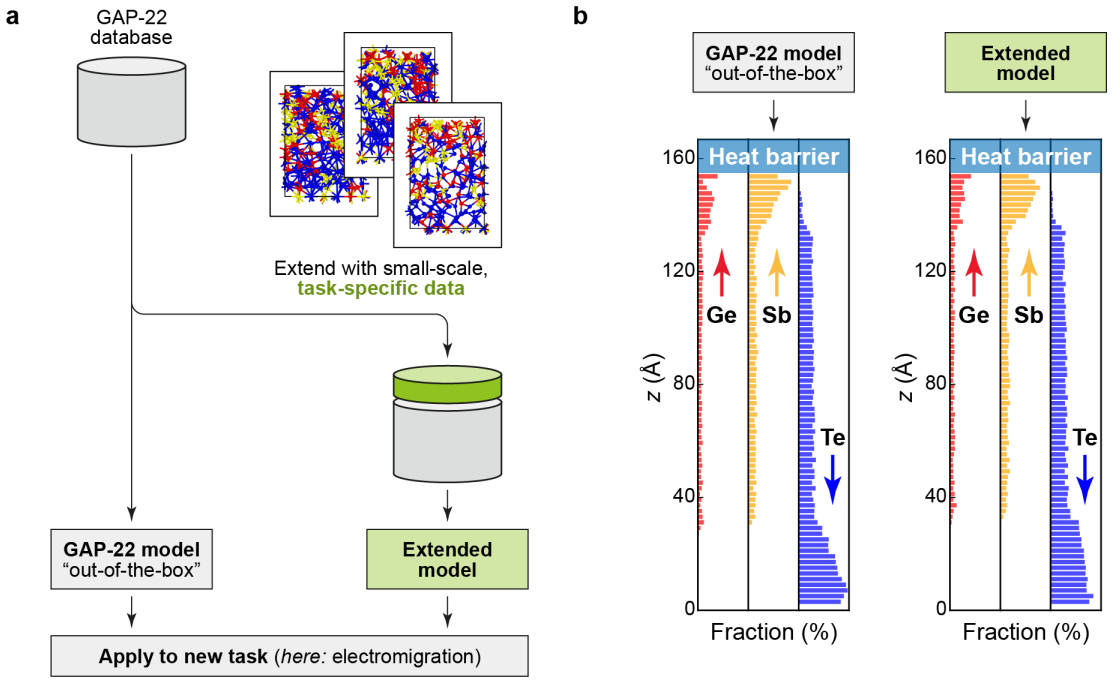


Figure 5.21: Extension of the GST-GAP-22 database to a new, task-specific application – Electromigration. (a) New configurations, simulated under external electric fields, were labelled with DFT and added to the original reference database, and a new GAP model was fitted based on this extended database. Both the GAP-22 model and the extended model can be used to study the electromigration process of GST. (b) The calculated elemental distribution obtained from both models: the left-hand side shows the data obtained with GST-GAP-22, also included in Fig. 5.22; the right-hand side shows the results of a simulation with the extended potential model. The fact that both results are very similar suggests that our current model can be used "out of the box" for simulations of electromigration in liquid GST.

To extend the database for the specific application case of electromigration, rock-salt-like structures of GST were set up, as done in the "domain-specific" iterations (viz. iter2-1, iter2-2, and iter2-3). The structural models were first melted and cooled to 1,000 K to obtain liquid phases. These structures were then annealed above 1,000 K, and electric fields were applied from the bottom to the top of the simulation cells, such as to trigger the electromigration of cation-like species (Ge and Sb) and anion-like ones (Te). Various electric fields with different magnitudes were used. One or two atomic layers were fixed (as barriers to stop continued atomic migration along the z direction that would be due to periodic boundary conditions) in most structural models; some were included where no atoms were fixed (to mimic the electromigration in the cell where no fixed boundary is met). The new dataset

contains additional 224 simulation cells (58,179 atoms); it was then labelled using PBEsol and added to the previous reference database. Another 1,200 representative ("sparse") points were used for those configurations in the new GAP model fit, in addition to the > 8,000 points in GST-GAP-22 (cf. Table 5.1).

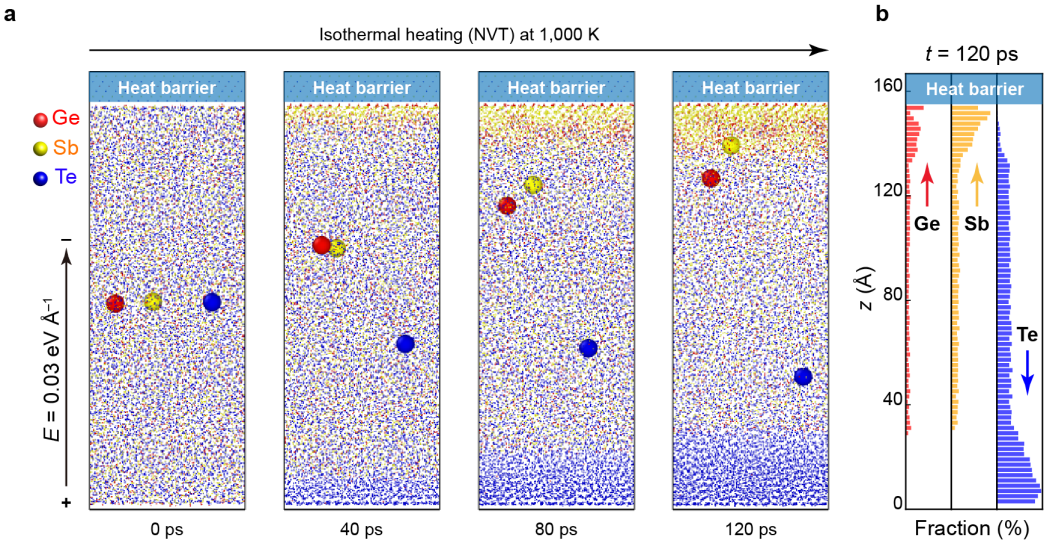


Figure 5.22: Electromigration and segregation in an electric field. (a) Simulation combining the GST-GAP-22 ML model and an applied electric field. An elongated structural model of $Ge_1Sb_2Te_4$ ($6 \times 6 \times 16.7$ nm^3, containing 18,816 atoms) was isothermally annealed at 1,000 K, i.e., in the liquid state corresponding to the beginning of the RESET process in devices. An electric field of 0.03 eV Å^{-1} was applied in z-direction, from the bottom to the top of the cell, as indicated by an arrow. The atomic charges are taken from Ref. 358. One atom of each species is highlighted by large spheres to exemplify the atomic segregation under applied electric field. (b) Distribution of the respective atomic species after 60,000 simulation steps.

With an extended GAP potential, the inclusion of electric field effects was considered in simulations, which is related to electromigration in the liquid state, and ultimately to device failure. Since the cumulated transient electric fields upon extensive cycling trigger significant electromigration over more than millions of SET and RESET operations in GST devices, reproducing the entire process directly is not feasible. Nonetheless, as a proof-of-concept, such electromigration processes can be accelerated using a large electric field of 0.03 eV Å^{-1}, applied from the bottom to the top of the cell, at 1,000 K. The resulting ML-driven MD trajectory, visualised in Fig. 5.22, shows a clear migration of cation- (anion-) like ions to the negative (positive) side, respectively. Notably, two GAP models, including the original GST-

GAP-22 potential and the extended one with the extended database (Fig. 5.21a), showed very similar results in the simulations of electromigration (Fig. 5.21b). This suggested the considerable transferability of the GST-GAP-22 potential to be used in some unseen structures. It is expected that such combination of ML potentials with electric field effects will play a central role in realistic modelling of PCM devices in the future.

5.6 Conclusion

As shown in this chapter, machine-learning potentials have enabled fully atomistic simulations of GST under realistic device geometry and programming conditions. Due to the much improved computational efficiency, the ML-driven simulations can be extended to even larger and more complex device geometries, in addition to longer timescales than in the present capability demonstrations. Beyond melting and crystal-growth simulations that are currently common in the field, it is expected that our ML model will allow for a thorough sampling of nucleation and the atomic-scale observation of the formation of grain boundaries in large models of GST, either under isothermal conditions or with a temperature gradient. In combination with advanced sampling techniques, ML-driven simulations should therefore allow to determine the nucleation barrier and critical nucleus size for GST. Moving from fundamental questions to those that are relevant to device engineering, in future work the inclusion of interface effects on the surrounding electrode and dielectric layers can be considered. For example, it has been reported that heat loss can be largely reduced by confining the PCM cell with Al_2O_3 walls, but their direct (atomic-scale) impact on thermal vibrations at the interface and on the phase-transition capacity of PCMs cannot be studied by FEM simulations alone. Atomistic ML models with extended reference databases are poised to provide this missing information, and in the future might help to predict the trend in minimum RESET energy as well as crystallisation time for different device geometries and the anticipation of microscopic failure mechanisms, enabling better architecture design. The

computationally guided design not just of individual PCM phases, but of the entire devices based on them, has therefore come within reach.

Chapter 6

Unlocking full-loop device-scale simulations

Following the demonstration of the device-scale MLMD simulation (containing more than half a million atoms) using a GAP potential developed for GST (as shown in Chapter 5), in this chapter, full-loop device-scale simulations with long time scale comparable to real device operations will be explored using new ML potentials. Note that this chapter includes some unpublished data. I would like to acknowledge the great contributions of Daniel F. Thomas du Toit for carrying out the hyperparameter optimisation and fitting the ACE models. I would also like to thank Stephen R. Elliott, Wei Zhang, and Volker L. Deringer for their guidance and useful discussions in this work.

6.1 Towards full-loop device simulations

Full-loop simulations (including both the RESET and SET processes) of GST-based devices are necessary, which lead to more comprehensive understanding of the programming mechanisms in GST-based devices. In contrast to RESET operations that could be completed using a relatively short pulse [357], SET operations require longer pulses (up to tens of nanoseconds), showing the stochastic nature of nucleation and growth in PCMs. Although GAP can be used in the demonstration of a RESET process of a device-scale model over 50 picoseconds (Fig. 5.18), or in a SET simulation using 12,096 atoms for several nanoseconds (Fig. 5.12), GAP-based MD simulations cannot simultaneously guarantee both large length scale (e.g., millions of atoms) and long time scale (e.g., several nanoseconds). Hence, given limited computational budgets, a balance between the simulation length scale and time scale should be considered when performing device-scale full-loop simulations using GAP

potentials.

A direct way to deal with this dilemma is to fit a new ML potential using a computationally more efficient framework, e.g., the atomic cluster expansion (ACE) as discussed in detail in Section 3.1. A large increase with a factor of at least 100 is expected for an ACE potential as compared to a GAP potential using the same reference database [182]. However, ACE potentials are less data-efficient than GAP, namely, ACE potentials require more training data. Collecting a new training database for a new ACE potential seems to be unrealistic and not productive. Alternatively, given a very diverse reference database used in the GAP potential as shown in Chapter 5 and showcased examples of how this training database can be extended to satisfy different simulation purposes, an extension can also be performed for the development of a new ACE. This is similar to the fitting process of ACE potentials for Te, as discussed in Section 4.3.1.

6.2 Unlocking billion-atom simulations

The strategy for fitting an ultrafast potential is based on the chemically and structurally diverse reference database of GST-GAP-22 [180], combined with largely improved computational efficiency from ACE [173, 182]. However, two major challenges remain: on one hand, hyperparameters used in ACE are difficult to manually tune throughout the fitting process; on the other hand, a physically reasonable ACE model requires much more training data to be fitted than GAP. Note that both difficulties stem from the mathematically complex form of ACE [172].

Starting from the original GST-GAP-22 training database (as mentioned in Chapter 5), a new potential was directly refitted using the ACE framework, in which hyperparameter optimisation was carried out using XPOT. However, the directly refitted and optimised ACE potential cannot well deal with the complex structural space of GST, giving some unphysical configurations during the MLMD simulations. For example, some atoms were clustering in the melt-quench simulations, leading to severe phase segregation in the amorphous structures. Note that these failure cases

are frequently observed at the early stage of developing ML potentials, due to the lack of training data.

To generate more training data for the ACE model, two strategies were implemented: on one hand, the two-step iterative training protocol was performed (including the standard and domain-specific iterations), as discussed in Chapter 5, which was combined with the addition of randomly sampled small-scale structures (containing tens of atoms); on the other hand, ACE-based random structure search (ACE-RSS) processes were carried out, similar with the GAP-RSS process mentioned in Section 3.2.2 and 4.3.2, which helped to sample a wider range of structural space of GST that was not sampled by the first strategy.

As shown in Fig. 6.1, starting from the initial ACE model, five iterations (iter-1 to iter-5) were carried out to generate new training data for the ACE potential. Firstly, an initial ACE model was fitted (iter-0) based on the original GST-GAP-22 database using the XPOT package, in which the space of hyperparameters can be automatically explored and optimised with minimal user input [207].

Note that the dimer data from the original GST-GAP-22 database was removed from the training database, similar to the fitting of the ACE potential for Te (cf. Section 4.3.2), because removing the dimers increased prediction accuracy on the testing dataset. In fact, these dimer data were included in GAP models so that the necessary two-body repulsive terms could be learned via the two-body structural descriptor in GAP. However, there is no clearly defined two-body term in the ACE model, and including some random structures (e.g., via ACE-RSS) would be a better option to capture the two-body repulsive term in ACE.

With the iter-0, three iterations (iter-1 to iter-3) were carried out, following a similar way as introduced in the previous work [180]: each iteration contained both the "standard"-type (i.e., conventional melt-quenched disordered structures) and the "domain-specific"-type configurations (i.e., intermediate configurations during the melting and crystallisation processes). However, due to the absence of specific

two-body term in ACE, the random structures that were created by the ab initio random structural search (AIRSS) code were manually added into each iteration, to avoid unphysical atomic clustering in the iterative structural models [183]. After iter-3, another XPOT optimisation was performed to ensure that the ACE model was well parametrised for the new iterative data. To further eliminate the unphysical configurations in ACE-driven molecular dynamics (ACE-MD) simulations, ACE-driven random structural search iterations (ACE-RSS) were carried out for two other iterations (iter-4 and iter-5), which included a wide range of structural models that were not sampled in iter-1 to iter-3. Note that such explorations of structural space have also been used in previous GAP-driven RSS (GAP-RSS) processes for different GAP potentials [178, 197, 201].

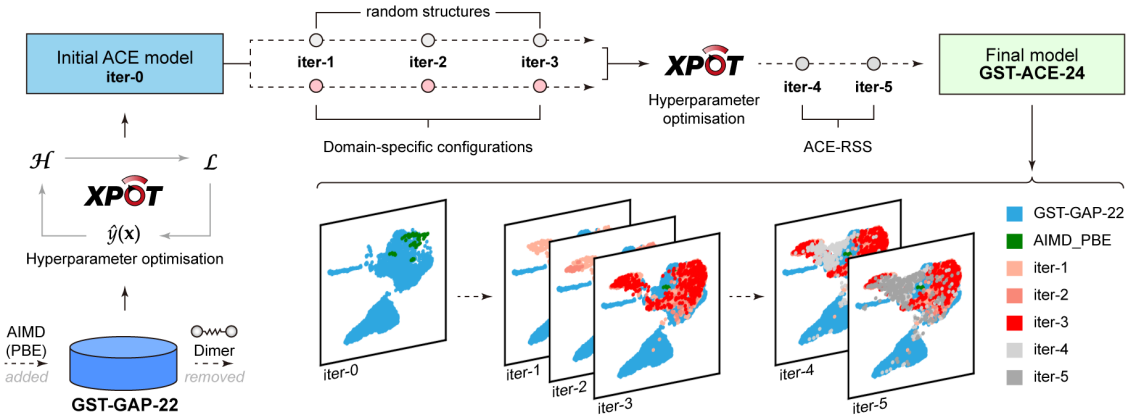


Figure 6.1: Fitting protocol for a new ACE potential for GST. Starting from the reference of GST-GAP-22, dimer data were all removed and five continuous iterations were carried out. The first three iterations (iter-1 to iter-3) include melt-quenched structures, intermediate configurations during the crystallisation and melting, and random structures generated via AIRSS code [198, 199]. The last two iterations (iter-4 and iter-5) contains the random structures generated via ACE-driven random structure search (ACE-RSS) processes. The resultant reference database is called "GST-ACE-24". The hyperparameter optimisations were carried out for twice: one on the GST-GAP-22 reference database, which provided a directly refitted ACE potential (the initial model); and the other one on the end of iter-3, which provided a well parametrised model for iter-4 and iter-5. The structural diversity of the reference database was shown during iterative training. Training structures in different iterations are visualised by UMAP embedding [359] of a kernel-based SOAP structural similarity metric [170].

The final training database (referred to as "GST-ACE-24" in the following) included 1,885 new training structures. All structures were labelled at the PBE level.

Only per-structure energy and per-atom forces were used in the fitting of the ACE potential, whereas per-structure stresses were not included. Another hyperparameter optimisation was carried out on the GST-ACE-24 reference database to obtain the final ACE potential.

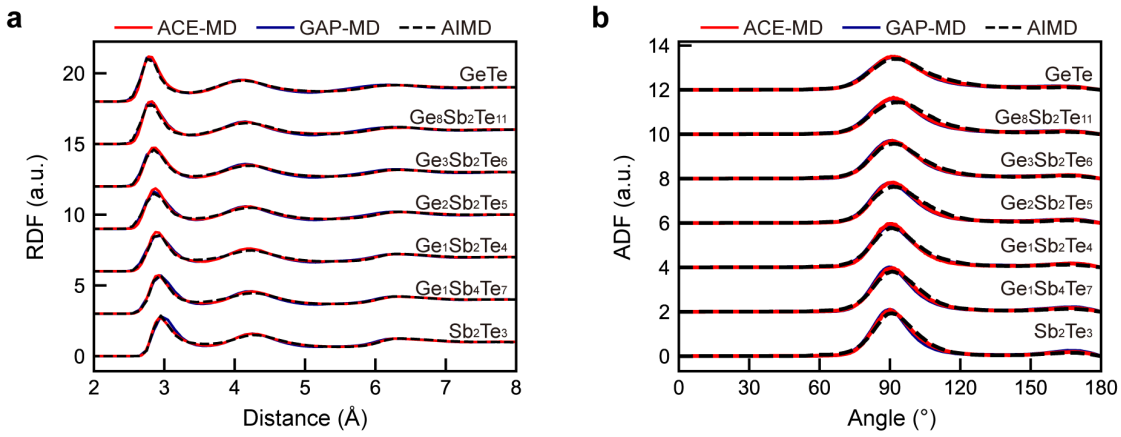


Figure 6.2: Structural properties of amorphous GST alloys obtained from ACE-MD (red), GAP-MD (blue), and AIMD (black) simulations. (a) The computed radial distribution function (RDF). (b) The calculated angular distribution function (ADF).

To validate the GST-ACE-24 potential, MLMD simulations were carried out to compare to the results from AIMD. Comprehensive structural analyses were carried out. The results of radial distribution function (RDF) and angular distribution function (ADF) were first obtained from the melt-quenched amorphous structures of seven different compositions using AIMD, GAP-MD, and ACE-MD. As shown in Fig. 6.2, the ACE-MD results agree well with the AIMD reference. Notably, ACE-MD outperformed the GAP-MD in some structural features compared to AIMD, indicating a good ability of the GST-ACE-24 model in the simulations of complex structures of GST alloys.

To demonstrate a largely increased computational efficiency, a direct comparison was carried out between GAP-MD and ACE-MD. The MD simulations were performed on the ARCHER2 high-performance cluster in the United Kingdom with the same MD settings. As shown in Fig. 6.3a, a weak scaling plot was generated by varying both the system size and the number of nodes proportionally, ensuring

that the computational load per node remained constant (i.e., 10,000 atoms per node). Good weak scaling behaviours were observed in both GAP and ACE models, which indicated that the computational workload is efficiently distributed across the nodes, with minimal overhead as the problem size grows. More importantly, the computational efficiency of ACE is almost ≈400 times larger than that of GAP, regardless of the number of nodes used in the simulations. In addition, different structural models of varied sizes (ranging from a thousand atoms to a billion atoms) were tested using GAP- and ACE-MD simulations (Fig. 6.3b). Different numbers of CPU cores were used, ranging from 128 CPU cores (corresponding to one node on ARCHER2) to 8192 CPU cores (corresponding to 64 nodes on ARCHER2). The MD steps per day were then recorded from each MD simulation.

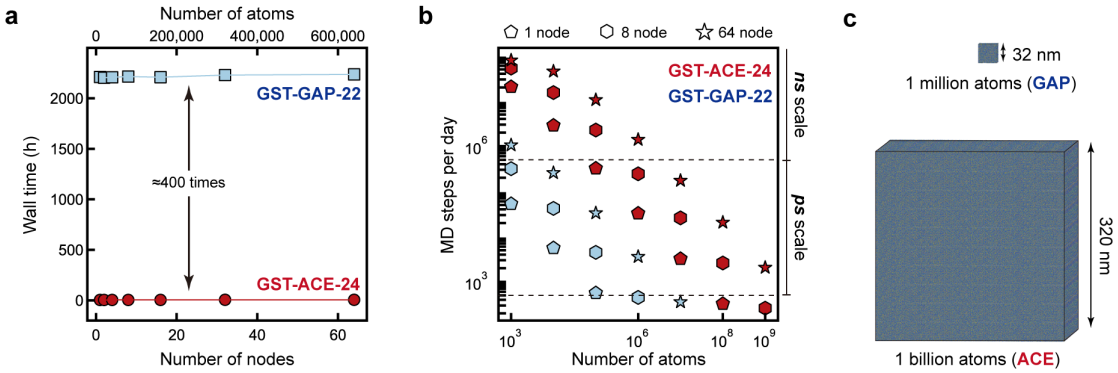


Figure 6.3: The comparison of computational efficiency between GAP-MD and ACE-MD. (a) The weak scaling plot of both GST-GAP-22 and GST-ACE-24 potentials. (b) The estimated MD steps per day as a function of system sizes. Different number of CPU cores on ARCHER2 were used to test the scaling capacity. (c) A direct comparison between ACE-MD and GAP-MD. Aiming at the same MD steps per day, ACE-MD can handle the system of around 1 billion atoms whereas GAP-MD can only be used for the structural model of several million atoms.

Based on the tested computational efficiencies, it is evident that with increased simulation sizes, the MD steps per day largely decreased, in a near-linear manner (Fig. 6.3b). Notably, adding more CPU cores can lead to an increased computational efficiency, suggesting good scaling capacity. A large efficiency gap was observed between ACE-MD and GAP-MD – ACE-MD is much faster than GAP-MD , achieving a speed-up of approximately two orders of magnitude. Moreover, GAP-MD simu-

lations require larger memory sizes, and simulations of millions of atoms failed due to the lack of memory. In contrast, ACE-MD did not encounter any memory issue and can still be used in exceedingly large-scale simulations up to one billion atoms.

By plotting the MD steps and system sizes against the simulation time scale and length scale (Fig. 6.3c), it is found that ACE-MD can satisfy the required computational efficiency for full-loop operations of GST-based devices, that is, device-scale MD simulations of millions of atoms (e.g., 0.5 to 100 million atoms) over tens of nanoseconds. Such superb computational efficiency also suggests a good application of ACE potentials in the study of complex phenomena in other material systems.

6.3 Simulating full-loop operations of crossbar devices

Based on the newly developed GST-ACE-24, ACE-MD simulations of RESET and SET operations in a crossbar device were carried out. In this chapter, the canonical ensemble (NVT, constant number of particles, volume, and temperature) and the microcanonical ensemble (NVE, constant number of particles, volume, and energy) were used. A Langevin thermostat (using the default damping value in LAMMPS) was used to control the temperature in the NVT simulations. The simulation of non-isothermal heating processes was performed using NVE simulations. Additional energy was added to the kinetic energy of the atoms in the programming regions, with a timestep of 2 ps. The timestep for all MLMD simulations was 2 fs. Figure 6.4 shows the 50-ps RESET process using ACE-MD in a structural model of 532,980 atoms, which was then compared to the GAP-MD result (cf. Fig. 5.18).

A 10-ps heating pulse was added along the vertical direction from the bottom to the top of the structural model. The simulations followed the same protocol discussed in Section 5.4.3, in which per-atom kinetic energies were added to the atoms to simulate the heating pulse. A ≈8 Å-thick slab was selected during the MD simulations, which served as a heat barrier and blocked the migration of atoms

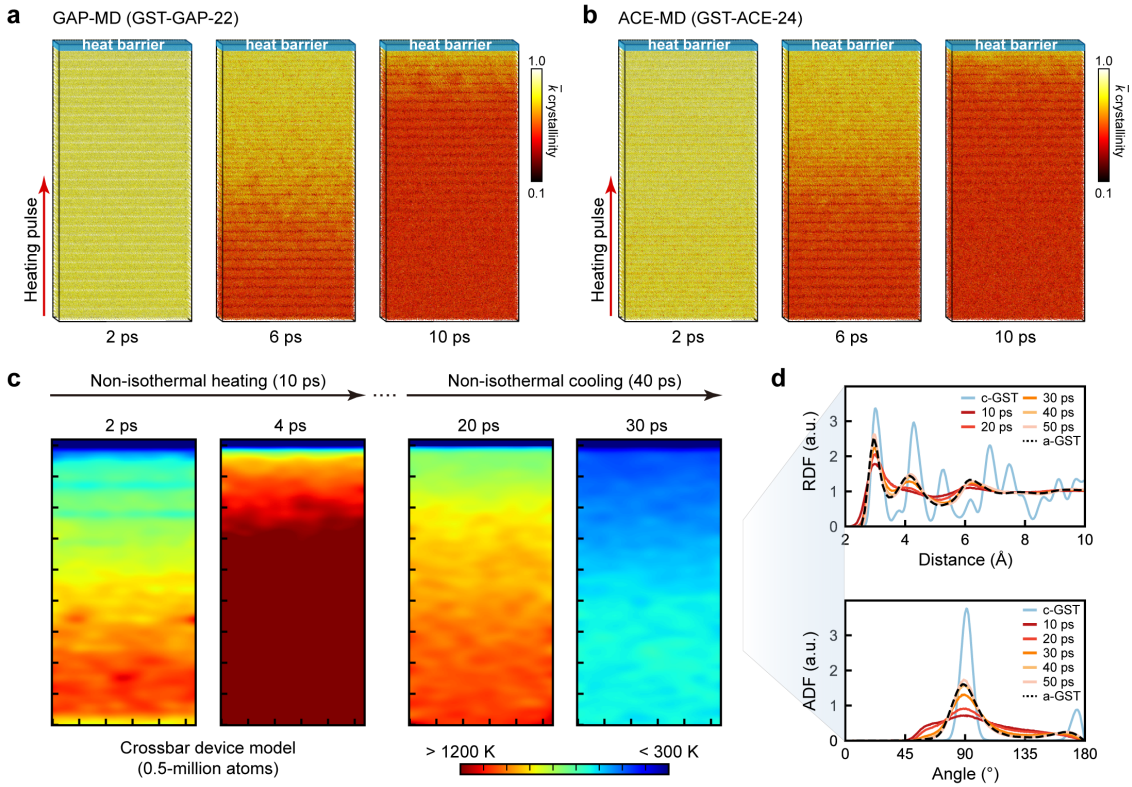


Figure 6.4: The RESET simulation of the crossbar device using ACE-MD. (a) The 10-ps heating simulation using GAP-MD as a reference. This panel was also shown in Fig. 5.18 [180]. (b) The 10-ps heating simulation using ACE-MD. Similar structural disordering was observed as compared to the GAP-MD reference. (c) The temperature evolution of the 10-ps heating pulse and the subsequent 40-ps cooling process, which was obtained using the same methods described in Section 5.4.3 and Fig. 5.18. (d) The computed radial distribution function (RDF) and angular distribution function (ADF) for the central area of the model (central 20 nm) during the RESET process. The results from small cells (cf. Fig. 5.12; "Bulk") were taken as a comparison.

across the periodic box. This slab was taken from a melt-quenched GST structure, and the atoms inside the slab were all fixed throughout the MD simulations. Similar melting behaviours were observed in both the ACE-MD and GAP-MD results, in which the structural disorder started from the bottom and gradually ran vertically across the entire crossbar model (Fig. 6.4a and b). Next, a 40-ps cooling process was simulated by gradually removing the per-atom kinetic energies from the atoms, as indicated by the computed temperature evolution during the cooling process (Fig. 6.4c). Based on the ACE-MD simulations, atomic information was obtained, showing the formation of amorphous GST after the 50-ps RESET process (Fig.

6.4d).

In addition to the RESET process, the SET process of the crossbar device can now be simulated using ACE-MD. Given that the width of the SET pulse (several nanoseconds) is much longer than that of the RESET pulse (10 ps). Hence, the SET process can be seen as an isothermal heating and thus an NVT ensemble was used. For nucleation-driven GST alloys, crystallisation can be separated into two steps: the homogeneous nucleation and the subsequent crystal growth (Fig. 6.5a).

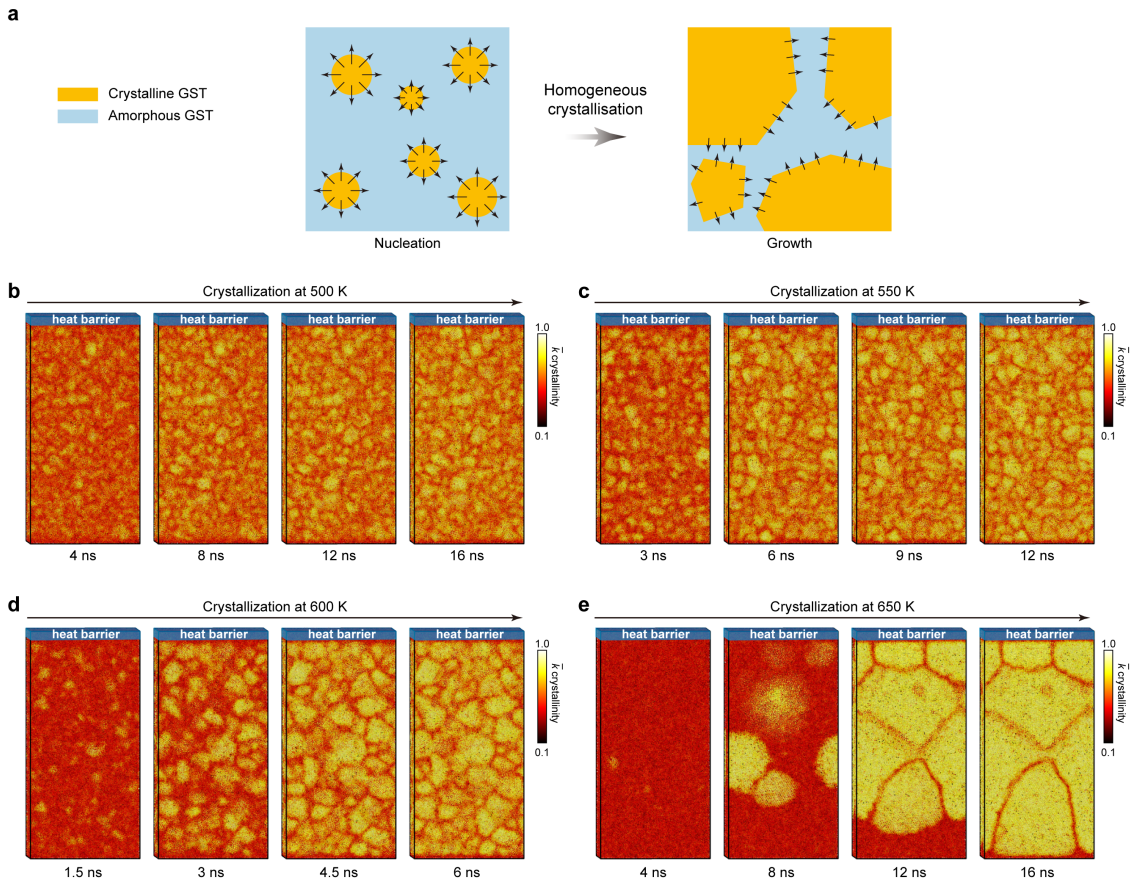


Figure 6.5: The SET simulation of the crossbar device using ACE-MD. (a) The sketch showing the crystallisation process of nucleation-driven GST alloys, which can be separated into two steps, i.e., homogeneous nucleation and crystal growth. The crystallisation of the cross-bar models was carried out at four different temperatures, including (b) 500 K, (c) 550 K, (d) 600 K, and (e) 650 K.

Figure 6.5b-e show the crystallisation of the structural model of crossbar devices at different temperatures. It was found that at lower temperatures, e.g., 500 and 550 K, abundant crystalline seeds appeared, indicating a fast nucleation process (Fig.

6.5b and c). These crystalline seeds grew rapidly, leading to the formation of polycrystal structures. When the annealing temperature increased to 600 K, the number of crystalline seeds was reduced (Fig. 6.5d). These nuclei quickly grew at 600 K, and formed a polycrystal SET structure with smaller grains but larger grain sizes as compared to the SET states at 500 and 550 K. In stark contrast, when crystallising at 650 K, no evident crystalline seeds were found until 4 ns, likely due to a largely reduced nucleation rate at 650 K (Fig. 6.5e). These surviving crystalline seeds then quickly grew, forming much larger grains in comparison to the SET structures at lower temperatures. Hence, a temperature-dependent crystallisation was observed in GST: at lower temperatures (e.g., 500 - 600 K), fast nucleation was observed, leading to a polycrystal-like SET structure with different grain morphology; in contrast, at higher temperatures (e.g., above 650 K), nucleation was slower but the growth was fast, resulting in fewer nuclei and thus fewer grains in the SET state. The temperature-dependent nucleation and growth were also supported by the experimentally measured results [360] and MLMD simulations on $Ge_2Sb_2Te_5$ [140] and Sb–Te alloys [361].

To quantify different grain morphology among the SET structures at different temperatures, the polyhedral template matching (PTM) method and the grain segmentation (GS) method were employed to distinguish different grains with different orientations in the recrystallised SET structures. Both methods are implemented in the Ovito software [258]. As shown in Fig. 6.6a, more grains were formed at lower annealing temperatures (e.g., 500 - 600 K) whereas higher temperatures (e.g., 650 K) resulted in very few but large grains. Different crystallisation behaviours were also quantified from the computed potential energies and crystal-like atoms during the SET process (Fig. 6.6b, c). When annealing at 500 and 550 K, the energies quickly dropped at the starting, together with a fast increase of crystal-like atoms, indicating fast nucleation processes. However, no energy drop was observed at the starting 4 ns at 650 K, implying a very slow nucleation process. Note that the lin-

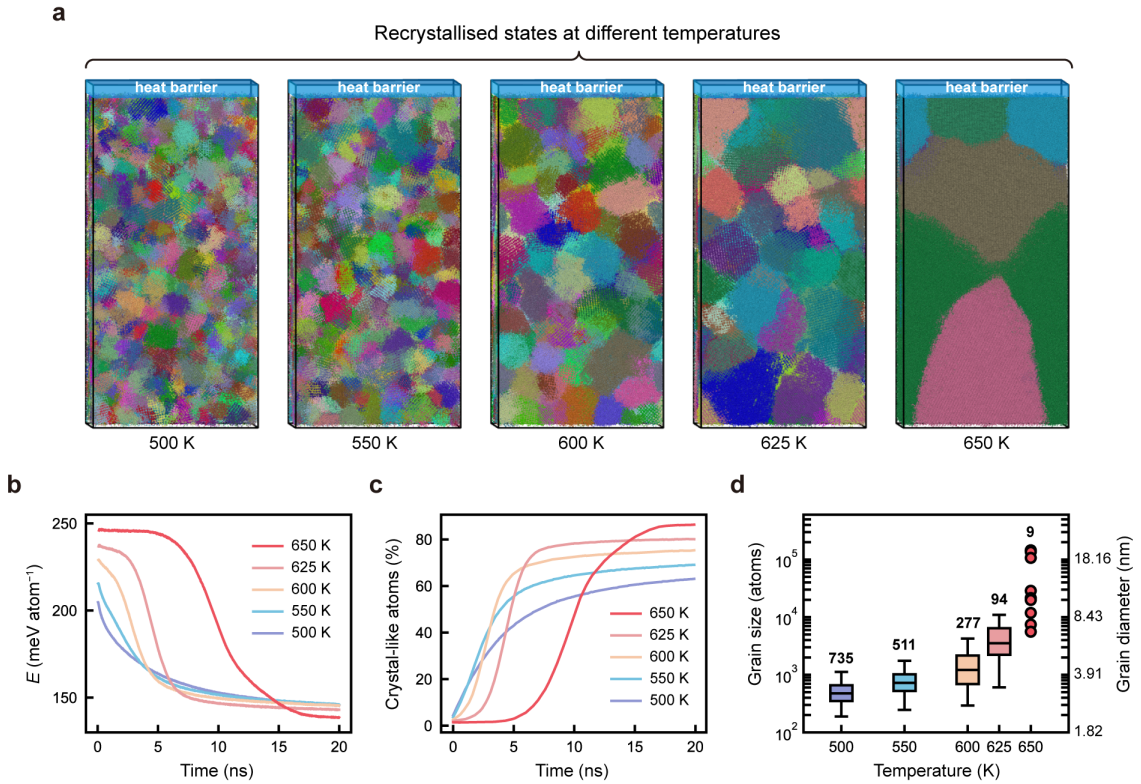


Figure 6.6: (a) Grain morphology of four different SET states at 500 K, 550 K, 600 K, 625 K, and 650 K. Different grains with different orientations indicated by different colours. (b) The computed potential energy during the crystallisation at different temperatures. (c) The computed crystal-like atoms during the crystallisation at different temperatures. (d) A box plot showing the grain-size distribution of different recrystallised structural models at 5 different temperatures. By assuming spherical grain shapes, the grain diameter, d, can be estimated via $d = \sqrt[3]{(6N)/(\pi\rho)}$, where N is the number of atoms in the grain and ρ is the atomic density of crystalline $Ge_1Sb_2Te_4$ (0.0318 $Å^{-3}$). In the box plot, the boxes indicate the range of grain sizes from 25% to 75% with a horizontal line indicating the median. The whiskers represent the full range of the data. The numbers above each box give the total number of grains.

ear decrease of the energy corresponds to a growth process. Hence, a fast growth was found at 600 and 650 K. In addition, the distributions of grain sizes were also computed (Fig. 6.6 d). A very rapid increase of grain sizes was observed for crystallisation above 600 K. It is found that nanosecond-scale recrystallisation below 625 K usually resulted in grain diameters of ≈2–8 nm, whereas high-temperature crystallisation (e.g., at 650 K) greatly increases the grain size to ≈5–20 nm.

6.4 Simulating full-loop operations of mushroom-type devices

Lastly, full-loop simulations (including both RESET and SET processes) of the mushroom-type device were demonstrated. In contrast to the volume-minimized confined cells in the crossbar array, the mushroom-type device adopts a contact-minimised geometry, indicating a larger volume of PCMs. Hence, the simulation size of the mushroom-type cell is larger than the confined cells in the crossbar array, requiring more computational resource.

In a recently reported 64-core mixed-signal in-memory computing chip, PCMs were used in a mushroom-type cell [50]. Based on the scanning transmission electron microscope (STEM) image, the size of the PCM cell was estimated to be 200 × 80 × 200 nm^3 , corresponding to a structural model of ≈100 million atoms.

Here a small-scale demonstration was provided using a two-dimensional slab model of 100 × 40 × 5 nm^3, which contained ≈800,000 atoms. This slab model can represent the cross-section in the centre of the original mushroom-type structural model (Fig. 6.7a). Note that similar two-dimensional slab models were also used in the previous FEM simulations to model the crystallisation of $Ge_2Sb_2Te_5$ [362], but in this thesis, the two-dimensional model with a decreased size as compared to the real mushroom-type device was used to reduce computational costs and achieve more efficient modelling. A full device-size modelling (containing more than 100 million atoms) will be shown in the future work.

A RESET operation using a 100-ps heating pulse was first simulated. The heat was only added in the programming region with a radius of 25 nm (indicated by the dashed semicircle in Fig. 6.7) by gradually increasing the per-atom kinetic energies of the atoms in this region (i.e., in a non-isothermal heating way). The modelling of this non-isothermal heating process was performed in the same way as discussed in Section 5.4.2. As shown in Fig. 6.7, the SOAP-based crystallinity similarity was

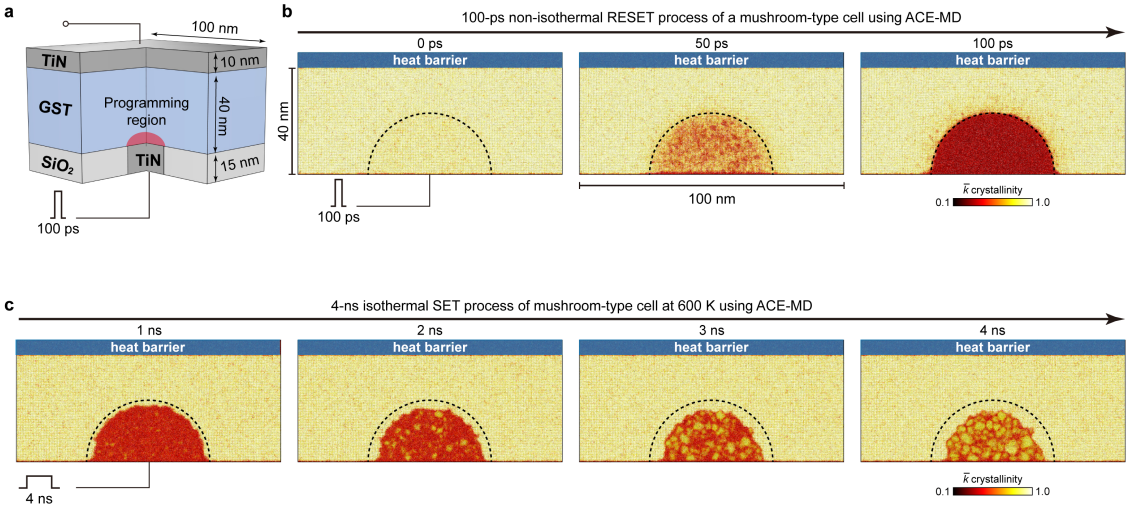


Figure 6.7: A demonstration of full-loop device-scale simulations of GST-based mushroom-type cell. (a) A three-dimensional sketch of mushroom-type cell. To model this cell, a near-two-dimensional slab structure was constructed (in a dimension of 100 $\times$ 40 $\times$ 5 nm^3; containing $\approx$800,000 atoms). A 100 $\times$ 5 $\times$ 5 nm^3 slab of atoms was fixed on the top (shallow blue), serving as a heat barrier to block atomic migration across the periodic box. (b) A RESET process of the mushroom-type structural model upon a 100-ps heating pulse. (c) A 4-ns SET process of the mushroom-type structural model. The exact programming region is indicated by the dashed black semicircle.

used to quantify the structural disordering process. The initial structural model is cubic $Ge_2Sb_2Te_5$. Upon non-isothermal heating, structural disordering appeared inside the programming region. Multiple melting centres were found, suggesting a collective melting behaviour inside the programming region. The disordering rapidly expanded across the programming region within 100 ps, leading to the formation of the amorphous area which was surrounded by the neighbour crystalline matrix.

After the non-isothermal RESET process, a 4-ns SET pulse was modelled using ACE-MD. Note that the width of the SET pulse (4 ns) is much longer than that of the RESET pulse (100 ps). Hence, the SET process can be seen as an isothermal heating and an NVT ensemble was used. At the starting of the SET process (before 1 ns), no crystalline nuclei were observed. However, the crystal growth proceeded at the crystalline-amorphous interface (i.e., the boundary of the programming region). As the crystal growth went on, some crystalline seeds appeared inside the amorphous area at 2 ns, showing simultaneously the crystal growth and nucleation. Upon

nucleation, these crystalline nuclei quickly grew up, forming polycrystal-like area inside the programming region. Note that the crystal growth proceeded continuously at the crystalline-amorphous interface throughout the SET process, and did not stop until encountering the growing crystalline seeds. In the resultant SET state, the size of the polycrystal-like area was smaller than that of the original programming region.

6.5 Conclusion

In this chapter, a domain-knowledge-based protocol was used to extend the reference database of the previous GST-GAP-22, including new structures from the standard iterations, domain-specific iterations, and ACE-RSS processes. Similar with the ACE potential developed for Te (as discussed in Chapter 5), the hyperparameters of the ACE models were optimised using XPOT (as introduced in Section 3.3). The ACE potential has been validated against the previous GAP potential and the AIMD reference, showing great accuracy in describing disordered structures and complex phase transition processes.

Based on the accurate and efficient ACE potential, the full-loop (including SET and RESET processes) (near-)device-scale simulations under realistic working conditions were demonstrated using ACE-MD, showing the capacity of ACE-MD in real-size device simulations. The structural models of both the confined cell in the crossbar device and the mushroom-type cell were constructed, containing hundreds of thousands of atoms. In the confined cell, a competition between the nucleation and crystal growth was observed: at lower temperatures (e.g., 500 - 600 K), the nucleation was dominant, leading to the formation of polycrystal SET states with abundant grains; at higher temperatures (e.g., above 650 K), the nucleation rate was largely decreased, and the fast crystal growth resulted in few and very large grains. The SET simulations of the mushroom-type cells showed simultaneous crystal growth and nucleation: the crystal growth proceeded continuously at the crystalline-amorphous interface, and in the meanwhile, nucleation occurred inside the amorphous region.

Chapter 7

Conclusion and outlook

7.1 Conclusion

This thesis has explored the application of machine-learned (ML) interatomic potentials to study the fundamental properties and switching mechanisms of phase-change materials (PCMs), focusing on elemental tellurium (Te), a phase-change selector material, and Ge–Sb–Te (GST) alloys, the flagship phase-change memory materials. A domain-knowledge-based iterative training protocol was proposed, and two fitting algorithms—the atomic cluster expansion (ACE) and Gaussian approximation potential (GAP)—were employed to develop ML potentials. The validations of ML potentials were carried out against density functional theory (DFT) and ab initio molecular dynamics (AIMD) references, demonstrating the ability to accurately simulate the structural properties and phase transitions of PCMs with significantly improved efficiency compared to AIMD. Large-scale ML-driven molecular dynamics (MLMD) simulations were performed, modelling the full-loop operations of PCM-based selector and memory devices under realistic conditions, at scales comparable to actual device geometries.

The thesis began with Te, an elemental system with significant structural complexity. To uncover the crystallisation mechanism of Te, bespoke GAP and ACE potentials were developed with optimised hyperparameters, enabling large-scale (102,000 atoms) MLMD simulations over tens of nanoseconds. It has been revealed that Te does not directly form its characteristic helical chains during nucleation. Instead, nucleation begins with the formation of disordered cubic-like motifs that rapidly transition into chiral helices within nanoseconds. Furthermore, the large-scale MLMD simulations uncovered the microscopic mechanism of the chirality transfer and the grain-boundary ordering during high-temperature anneal-

ing, mediated by disordered cubic-like motifs. These understandings were crucial for elucidating the switching mechanisms of Te-based selector devices: the volatile switching behaviour was attributed to the rapid and spontaneous crystallisation during cooling, and the large variation in switch-on voltages was linked to the degree of polycrystallinity in the switch-off states.

Next, the thesis extended to the more chemically and structurally complex GST alloys. A two-step iterative training protocol was used to create a diverse and comprehensive reference database, which included both standard iterations (containing melt-quenched disordered structures) and domain-specific iterations (containing intermediate configurations during crystallisation and melting). Based on this reference database, a GAP potential (referred to as "GST-GAP-22") was developed. After thorough validation against DFT and AIMD references, the GST-GAP-22 model enabled accurate and efficient simulations of GST structures and phase transitions. Different simulation protocols were proposed, including iterative RESET, cumulative SET, and non-isothermal heating, to mimic the realistic programming operations and complex working conditions in real-world PCM-based devices. To showcase the ability to model real-world GST-based devices, a 50-ps GAP-driven molecular dynamics (GAP-MD) simulation of the RESET process was demonstrated, using a device-scale structural model containing over half a million atoms.

To achieve full-loop device-scale simulations of GST-based devices, an ACE potential was developed with optimised hyperparameters and based on the extension of the GST-GAP-22 reference database through multiple domain-knowledge-based iterations. This ACE potential has provided much improved computational efficiency at least two orders of magnitude faster than GST-GAP-22. Such enhanced speed has allowed for simulations of nanosecond switching processes in device-scale models under realistic conditions, achieving full-loop atomistic modelling of GST-based devices. The ACE-driven molecular dynamics simulations revealed that at lower temperatures (e.g., 500–600 K), nucleation dominated the crystallisation process

in confined GST cells, resulting in polycrystalline SET structures with numerous small grains. In contrast, higher temperatures (e.g., above 650 K) promoted rapid crystal growth, leading to the formation of fewer, but much larger, crystal grains. This temperature-dependent competition between nucleation and crystal growth is crucial for understanding the microscopic switching mechanisms of GST-based devices. Additionally, near-device-scale MLMD simulations of mushroom-type cells over several nanoseconds revealed simultaneous nucleation and growth processes.

7.2 Outlook

The findings presented in this thesis emphasise the significant role of machine-learned interatomic potentials in advancing the study of PCMs. The ability to perform atomistic simulations at scales that closely resemble real-world device geometries and working conditions opens new avenues for exploring complex programming mechanisms and optimising the performance of practical devices.

However, challenges remain in simulating real GST-based devices. One key challenge is the need to account for complex interface effects. In actual PCM-based devices, the PCM layers are connected to other functional units, such as selectors or electrodes. Additionally, PCM bits are typically surrounded by dielectric materials (e.g., SiO_2). These interfaces might significantly influence the structural properties and switching behaviour of the adjacent PCMs. Modelling these complex interfaces requires large simulation sizes, highlighting the need for accurate and efficient ML potentials tailored to structurally and chemically more complex heterogeneous interface systems.

Another challenge is the accurate modelling of coupled thermal and electric conditions. As demonstrated in Chapters 5 and 6, heating pulses can be simulated using non-isothermal conditions, where additional per-atom kinetic energies are applied to the atoms in the programming region. However, in real-world PCM-based devices, the temperature gradients generated by Joule heating are more complex than those modelled in this thesis. Furthermore, Joule heating under an external electric field

might cause elemental migration upon extensive cycling. Instead of using constant per-atom charges, as demonstrated in Chapter 5 (Section 5.5.2), real atomic charges are highly dependent on local atomic environments. Therefore, the development of ML potentials incorporating charge information is necessary to better simulate electromigration in PCM-based devices.

Moreover, different simulation methods, as depicted in the simulation hierarchy in Fig. 1.21, can be integrated into device simulations. For example, thermoelectric effects in GST-based devices can be effectively described using finite element method (FEM) simulations, which provide accurate estimates of temperature distribution during the switching process. However, FEM simulations lack atomic-level detail and cannot capture the stochastic nature of SET and RESET operations. To address this, FEM simulations can be combined with MLMD simulations, where temperature data from FEM simulations are used in MLMD simulations, and atomic-level information from the phase-change processes can be fed back into the FEM models.

In conclusion, this thesis provides a comprehensive foundation for the use of machine-learned interatomic potentials in the study of PCMs, which paves the way for the modelling and innovation of next-generation PCM-based devices.

Bibliography

[1] Zhang, W., Mazzarello, R., Wuttig, M. & Ma, E. Designing crystallization in phase-change materials for universal memory and neuro-inspired computing. *Nat. Rev. Mater.* **4**, 150–168 (2019).

[2] Ovshinsky, S. R. Reversible electrical switching phenomena in disordered structures. *Phys. Rev. Lett.* **21**, 1450–1453 (1968).

[3] Chen, M., Rubin, K. A. & Barton, R. W. Compound materials for reversible, phase-change optical data storage. *Appl. Phys. Lett.* **49**, 502–504 (1986).

[4] Yamada, N. *et al.* High Speed Overwritable Phase Change Optical Disk Material. *Jpn. J. Appl. Phys.* **26**, 61 (1987).

[5] Lankhorst, M. H., Ketelaars, B. W. & Wolters, R. A. Low-cost and nanoscale non-volatile memory concept for future silicon chips. *Nat. Mater.* **4**, 347–352 (2005).

[6] Wuttig, M. *et al.* The role of vacancies and local distortions in the design of new phase-change materials. *Nat. Mater.* **6**, 122–128 (2007).

[7] Wuttig, M. Towards a universal memory. *Nat. Mater.* **4**, 265–266 (2005).

[8] Yoon, K. J., Kim, Y. & Hwang, C. S. What Will Come After V-NAND—Vertical Resistive Switching Memory? *Adv. Electron. Mater.* **5**, 1800914 (2019).

[9] Brennan, K. F. & Summers, C. J. Theory of resonant tunneling in a variably spaced multiquantum well structure: An Airy function approach. *J. Appl. Phys.* **61**, 614–623 (1987).

[10] Zhang, B. *et al.* Element-resolved atomic structure imaging of rocksalt $Ge_2Sb_2Te_5$ phase-change material. *Appl. Phys. Lett.* **108**, 191902 (2016).

[11] Zhang, W. *et al.* Role of vacancies in metal-insulator transitions of crystalline phase-change materials. *Nat. Mater.* **11**, 952–956 (2012).

[12] Yamada, N., Ohno, E., Nishiuchi, K., Akahira, N. & Takao, M. Rapid-phase transitions of GeTe–Sb_2Te_3 pseudobinary amorphous thin films for an optical disk memory. *J. Appl. Phys.* **69**, 2849–2856 (1991).

[13] Matsunaga, T. & Yamada, N. Structural investigation of $GeSb_2Te_4$: A high-speed phase-change material. *Phys. Rev. B* **69**, 104111 (2004).

[14] Kooi, B. J. & De Hosson, J. T. M. Electron diffraction and high-resolution transmission electron microscopy of the high temperature crystal structures of $Ge_xSb_2Te_{3+x}$ (x=1,2,3) phase change material. *J. Appl. Phys.* **92**, 3584–3590 (2002).

[15] Zhu, M. *et al.* Direct observation of partial disorder and zipperlike transition in crystalline phase change materials. *Phys. Rev. Mater.* **3**, 033603 (2019).

[16] Akola, J. *et al.* Experimentally constrained density-functional calculations of the amorphous structure of the prototypical phase-change material $Ge_2Sb_2Te_5$. *Phys. Rev. B* **80**, 020201 (2009).

[17] Kohara, S. *et al.* Structural basis for the fast phase change of $Ge_2Sb_2Te_5$: Ring statistics analogy between the crystal and amorphous states. *Appl. Phys. Lett.* **89**, 201910 (2006).

[18] Boniardi, M. *et al.* A physics-based model of electrical conduction decrease with time in amorphous $Ge_2Sb_2Te_5$. *J. Appl. Phys.* **105**, 084506 (2009).

[19] Sebastian, A. *et al.* Tutorial: Brain-inspired computing using phase-change memory devices. *J. Appl. Phys.* **124**, 111101 (2018).

[20] Fantini, P., Brazzelli, S., Cazzini, E. & Mani, A. Band gap widening with time induced by structural relaxation in amorphous $Ge_2Sb_2Te_5$ films. *Appl. Phys. Lett.* **100**, 013505 (2012).

[21] Raty, J. Y. *et al.* Aging mechanisms in amorphous phase-change materials. *Nat. Commun.* **6**, 7467 (2015).

[22] Luckas, J. *et al.* Defects in amorphous phase-change materials. *J. Mater. Res.* **28**, 1139–1147 (2013).

[23] Luckas, J. M. *Electronic Transport in Amorphous Phase-Change Materials.* Ph.D. thesis (2012).

[24] Zhu, M., Ren, K. & Song, Z. Ovonic threshold switching selectors for three-dimensional stackable phase-change memory. *MRS Bull.* **44**, 715–720 (2019).

[25] Wong, H.-S. P. *et al.* Phase change memory. *Proc. IEEE* **98**, 2201 (2010).

[26] Northover, W. R. & Pearson, A. D. Glass composition (1964).

[27] Ovshinsky, S. R. Symmetrical current controlling device (1966).

[28] Xu, M. *et al.* Pressure tunes electrical resistivity by four orders of magnitude in amorphous $Ge_2Sb_2Te_5$ phase-change memory alloy. *Proc. Natl. Acad. Sci.* **109**, E1055–E1062 (2012).

[29] Anbarasu, M., Wimmer, M., Bruns, G., Salinga, M. & Wuttig, M. Nanosecond threshold switching of $GeTe_6$ cells and their potential as selector devices. *Appl. Phys. Lett.* **100**, 143505 (2012).

[30] Manivannan, A., Myana, S. K., Miriyala, K., Sahu, S. & Ramadurai, R. Low power ovonic threshold switching characteristics of thin $GeTe_6$ films using conductive atomic force microscopy. *Appl. Phys. Lett.* **105**, 243501 (2014).

[31] Kuzum, D., Jeyasingh, R. G., Lee, B. & Wong, H. S. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Lett.* **12**, 2179–2186 (2012).

[32] Lee, M.-J. *et al.* A plasma-treated chalcogenide switch device for stackable scalable 3D nanoscale memory. *Nat. Commun.* **4**, 2629 (2013).

[33] Velea, A. *et al.* Te-based chalcogenide materials for selector applications. *Sci. Rep.* **7**, 8103 (2017).

[34] Yoo, J., Lee, D., Park, J., Song, J. & Hwang, H. Steep Slope Field-Effect Transistors With B–Te-Based Ovonic Threshold Switch Device. *IEEE J. Electron Devices Soc.* **6**, 821–824 (2018).

[35] Chekol, S. A. *et al.* A C–Te-based binary OTS device exhibiting excellent performance and high thermal stability for selector application. *Nanotechnology* **29**, 345202 (2018).

[36] Kim, S.-D. *et al.* Effect of Ge Concentration in Ge_xSe_{1-x} Chalcogenide Glass on the Electronic Structures and the Characteristics of Ovonic Threshold Switching (OTS) Devices. *ECS Solid State Lett.* **2**, Q75 (2013).

[37] Shin, S.-Y. *et al.* The effect of doping Sb on the electronic structure and the device characteristics of Ovonic Threshold Switches based on Ge-Se. *Sci. Rep.* **4**, 7099 (2014).

[38] Avasarala, N. S. *et al.* Half-threshold bias Ioff reduction down to nA range of thermally and electrically stable high-performance integrated OTS selector, obtained by Se enrichment and N-doping of thin GeSe layers. In *2018 IEEE Symposium on VLSI Technology*, 209–210 (2018).

[39] Avasarala, N. S. *et al.* Doped GeSe materials for selector applications. In *2017 47th European Solid-State Device Research Conference (ESSDERC)*, 168–171 (2017).

[40] Cheng, H. Y. *et al.* Ultra-High Endurance and Low IOFF Selector based on AsSeGe Chalcogenides for Wide Memory Window 3D Stackable Cross-point Memory. In *2018 IEEE International Electron Devices Meeting (IEDM)*, 37.3.1–37.3.4 (2018).

[41] Gallo, M. L. & Sebastian, A. An overview of phase-change memory device physics. *J. Phys. D: Appl. Phys.* **53**, 213002 (2020).

[42] Gong, N. *et al.* A no-verification multi-level-cell (MLC) operation in cross-point OTS-PCM. In *2020 IEEE Symposium on VLSI Technology*, 1–2 (2020).

[43] Fazio, A. Advanced technology and systems of cross point memory. In *2020 IEEE International Electron Devices Meeting (IEDM)*, 24.1.1–24.1.4 (2020).

[44] Kau, D. *et al.* A stackable cross point Phase Change Memory. In *2009 IEEE International Electron Devices Meeting (IEDM)*, 1–4 (2009).

[45] Cheng, H.-Y., Carta, F., Chien, W.-C., Lung, H.-L. & BrightSky, M. J. 3D cross-point phase-change memory for storage-class memory. *J. Phys. D: Appl. Phys.* **52**, 473002 (2019).

[46] Rao, F., Zhang, W. & Ma, E. Catching structural transitions in liquids. *Science* **364**, 1032–1033 (2019).

[47] Ding, K. *et al.* Phase-change heterostructure enables ultralow noise and drift for memory operation. *Science* **366**, 210–215 (2019).

[48] Boybat, I. *et al.* Neuromorphic computing with multi-memristive synapses. *Nat. Commun.* **9**, 2514 (2018).

[49] Le Gallo, M. *et al.* Mixed-precision in-memory computing. *Nat. Electron.* **1**, 246–253 (2018).

[50] Le Gallo, M. *et al.* A 64-core mixed-signal in-memory compute chip based on phase-change memory for deep neural network inference. *Nat. Electron.* **6**, 680–693 (2023).

[51] Ambrogio, S. *et al.* An analog-AI chip for energy-efficient speech recognition and transcription. *Nature* **620**, 768–775 (2023).

[52] Mead, C. Neuromorphic electronic systems. *Proceedings of the IEEE* **78**, 1629–1636 (1990).

[53] Merolla, P. A. *et al.* A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**, 668–673 (2014).

[54] Service, R. F. The brain chip. *Science* **345**, 614–616 (2014).

[55] Yang, J. J., Strukov, D. B. & Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13–24 (2013).

[56] Burr, G. W. *et al.* Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2**, 89–124 (2016).

[57] Zidan, M. A., Strachan, J. P. & Lu, W. D. The future of electronics based on memristive systems. *Nat. Electron.* **1**, 22–29 (2018).

[58] Nandakumar, S. R. *et al.* A phase-change memory model for neuromorphic computing. *J. Appl. Phys.* **124**, 152135 (2018).

[59] Wuttig, M., Bhaskaran, H. & Taubner, T. Phase-change materials for non-volatile photonic applications. *Nat. Photon.* **11**, 465–476 (2017).

[60] Wright, C. D., Bhaskaran, H. & Pernice, W. H. P. Integrated phase-change photonic devices and systems. *MRS Bull.* **44**, 721–727 (2019).

[61] Rios, C., Hosseini, P., Wright, C. D., Bhaskaran, H. & Pernice, W. H. P. On-Chip Photonic Memory Elements Employing Phase-Change Materials. *Adv. Mater.* **26**, 1372–1377 (2014).

[62] Ríos, C. *et al.* Integrated all-photonic non-volatile multi-level memory. *Nat. Photon.* **9**, 725–732 (2015).

[63] Wang, Q. *et al.* Optically reconfigurable metasurfaces and photonic devices based on phase change materials. *Nat. Photon.* **10**, 60–65 (2016).

[64] Kuramochi, E. & Notomi, M. Optical memory: Phase-change memory. *Nat. Photon.* **9**, 712–714 (2015).

[65] Cheng, Z. *et al.* Device-Level Photonic Memories and Logic Applications Using Phase-Change Materials. *Adv. Mater.* **30**, 1802435 (2018).

[66] Farmakidis, N., Dong, B. & Bhaskaran, H. Integrated photonic neuromorphic computing: Opportunities and challenges. *Nat. Rev. Electr. Eng.* **1**, 358–373 (2024).

[67] Zhou, W. *et al.* In-memory photonic dot-product engine with electrically programmable weight banks. *Nat. Commun.* **14**, 2887 (2023).

[68] Cheng, Z., Rios, C., Pernice, W. H. P., Wright, C. D. & Bhaskaran, H. On-chip photonic synapse. *Sci. Adv.* **3**, e1700160 (2017).

[69] Feldmann, J. *et al.* Calculating with light using a chip-scale all-optical abacus. *Nat. Commun.* **8**, 1256 (2017).

[70] Feldmann, J., Youngblood, N., Wright, C. D., Bhaskaran, H. & Pernice, W. H. P. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**, 208–214 (2019).

[71] Ríos, C. *et al.* In-memory computing on a photonic platform. *Sci. Adv.* **5**, eaau5759 (2019).

[72] Hosseini, P., Wright, C. D. & Bhaskaran, H. An optoelectronic framework enabled by low-dimensional phase-change films. *Nature* **511**, 206–211 (2014).

[73] Tao, S. *et al.* Phase Change Materials for Nonvolatile, Solid-State Reflective Displays: From New Structural Design Rules to Enhanced Color-Changing Performance. *Adv. Opt. Mater.* **8**, 2000062 (2020).

[74] Carrillo, S. G.-C. *et al.* A Nonvolatile Phase-Change Metamaterial Color Display. *Adv. Opt. Mater.* **7**, 1801782 (2019).

[75] Ríos, C., Hosseini, P., Taylor, R. A. & Bhaskaran, H. Color Depth Modulation and Resolution in Phase-Change Material Nanodisplays. *Adv. Mater.* **28**, 4720–4726 (2016).

[76] Siegrist, T. *et al.* Disorder-induced localization in crystalline phase-change materials. *Nat. Mater.* **10**, 202–208 (2011).

[77] Gabardi, S., Caravati, S., Sosso, G. C., Behler, J. & Bernasconi, M. Microscopic origin of resistance drift in the amorphous state of the phase-change compound GeTe. *Phys. Rev. B* **92**, 054201 (2015).

[78] Shportko, K. *et al.* Resonant bonding in crystalline phase-change materials. *Nat. Mater.* **7**, 653–658 (2008).

[79] Zhang, W., Mazzarello, R. & Ma, E. Phase-change materials in electronics and photonics. *MRS Bull.* **44**, 686–690 (2019).

[80] Fong, S. W., Neumann, C. M. & Wong, H. S. P. Phase-Change Memory—Towards a Storage-Class Memory. *IEEE Trans. Electron Devices* **64**, 4374–4385 (2017).

[81] Kim, C. *et al.* Fullerene thermal insulation for phase change memory. *Appl. Phys. Lett.* **92**, 013109 (2008).

[82] Ahn, C. *et al.* Energy-Efficient Phase-Change Memory with Graphene as a Thermal Barrier. *Nano Lett.* **15**, 6809–6814 (2015).

[83] Fong, S. W. *et al.* Dual-Layer Dielectric Stack for Thermally Isolated Low-Energy Phase-Change Memory. *IEEE Transactions on Electron Devices* **64**, 4496–4502 (2017).

[84] Jin, B., Kang, D., Kim, J., Meyyappan, M. & Lee, J.-S. Thermally efficient and highly scalable In_2Se_3 nanowire phase change memory. *J. Appl. Phys.* **113**, 164303 (2013).

[85] Reifenberg, J. P. *et al.* Thermal Boundary Resistance Measurements for Phase-Change Memory Devices. *IEEE Electron Device Lett.* **31**, 56–58 (2010).

[86] Lee, B. S. *et al.* Observation of the role of subcritical nuclei in crystallization of a glassy solid. *Science* **326**, 980–984 (2009).

[87] Salinga, M. *et al.* Measurement of crystal growth velocity in a melt-quenched phase-change material. *Nat. Commun.* **4**, 2371 (2013).

[88] Ielmini, D., Lacaita, A. L. & Mantegazza, D. Recovery and Drift Dynamics of Resistance and Threshold Voltages in Phase-Change Memories. *IEEE Trans. Electron Devices* **54**, 308–315 (2007).

[89] Caravati, S., Bernasconi, M., Kühne, T. D., Krack, M. & Parrinello, M. Co-existence of tetrahedral- and octahedral-like sites in amorphous phase change materials. *Appl. Phys. Lett.* **91**, 171906 (2007).

[90] Caravati, S., Bernasconi, M., Kühne, T. D., Krack, M. & Parrinello, M. First principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects. *J. Phys. Cond. Matt.* **21**, 255501 (2009).

[91] Akola, J. & Jones, R. O. Density functional study of amorphous, liquid and crystalline $Ge_2Sb_2Te_5$: Homopolar bonds and/or AB alternation? *J. Phys. Cond. Matt.* **20**, 465103 (2008).

[92] Akola, J. & Jones, R. O. Structure of liquid phase change material AgInSbTe from density functional/molecular dynamics simulations. *Appl. Phys. Lett.* **94**, 251905 (2009).

[93] Chen, Y. *et al.* Chemical understanding of resistance drift suppression in Ge–Sn–Te phase-change memory materials. *J. Mater. Chem. C* **8**, 71–77 (2020).

[94] Wełnic, W. *et al.* Unravelling the interplay of local structure and physical properties in phase-change materials. *Nat. Mater.* **5**, 56–62 (2006).

[95] Sun, Z., Zhou, J. & Ahuja, R. Unique melting behavior in phase-change materials for rewritable data storage. *Phys. Rev. Lett.* **98**, 055505 (2007).

[96] Friedrich, I., Weidenhof, V., Njoroge, W., Franz, P. & Wuttig, M. Structural transformations of $Ge_2Sb_2Te_5$ films studied by electrical resistance measurements. *J. Appl. Phys.* **87**, 4130–4134 (2000).

[97] Rosenthal, T., Schneider, M. N., Stiewe, C., Döblinger, M. & Oeckler, O. Real Structure and Thermoelectric Properties of GeTe-Rich Germanium Antimony Tellurides. *Chem. Mater.* **23**, 4349–4356 (2011).

[98] Zhang, B. *et al.* Vacancy Structures and Melting Behavior in Rock-Salt GeSbTe. *Sci. Rep.* **6**, 25453 (2016).

[99] Wang, X.-P. *et al.* Element-specific amorphization of vacancy-ordered GeSbTe for ternary-state phase change memory. *Acta Mater.* **136**, 242–248 (2017).

[100] Hegedüs, J. & Elliott, S. R. Microscopic origin of the fast crystallization ability of Ge-Sb-Te phase-change memory materials. *Nat. Mater.* **7**, 399–405 (2008).

[101] Lee, T. H. & Elliott, S. R. Ab Initio computer simulation of the early stages of crystallization: Application to $Ge_2Sb_2Te_5$ phase-change materials. *Phys. Rev. Lett.* **107**, 145702 (2011).

[102] Skelton, J. M., Pallipurath, A. R., Lee, T.-H. & Elliott, S. R. Atomistic Origin of the Enhanced Crystallization Speed and n-Type Conductivity in Bi-doped Ge-Sb-Te Phase-Change Materials. *Adv. Funct. Mater.* **24**, 7291–7300 (2014).

[103] Kalikka, J., Akola, J., Larrucea, J. & Jones, R. O. Nucleus-driven crystallization of amorphous $Ge_2Sb_2Te_5$: A density functional study. *Phys. Rev. B* **86**, 144113 (2012).

[104] Kalikka, J., Akola, J. & Jones, R. O. Simulation of crystallization in $Ge_2Sb_2Te_5$: A memory effect in the canonical phase-change material. *Phys. Rev. B* **90**, 184109 (2014).

[105] Kalikka, J., Akola, J. & Jones, R. O. Crystallization processes in the phase change material $Ge_2Sb_2Te_5$: Unbiased density functional/molecular dynamics simulations. *Phys. Rev. B* **94**, 134105 (2016).

[106] Branicio, P. S. *et al.* Atomistic insights into the nanosecond long amorphization and crystallization cycle of nanoscale $Ge_2Sb_2Te_5$: An ab initio molecular dynamics study. *Phys. Rev. Mater.* **2**, 043401 (2018).

[107] Bai, K., Tan, T. L., Branicio, P. S. & Sullivan, M. B. Time-temperature-transformation and continuous-heating-transformation diagrams of $GeSb_2Te_4$ from nanosecond-long *ab initio* molecular dynamics simulations. *Acta Mater.* **121**, 257–265 (2016).

[108] Loke, D. *et al.* Breaking the speed limits of phase-change memory. *Science* **336**, 1566–1569 (2012).

[109] Ronneberger, I., Zhang, W., Eshet, H. & Mazzarello, R. Crystallization properties of the $Ge_2Sb_2Te_5$ phase-change compound from advanced simulations. *Adv. Funct. Mater.* **25**, 6407–6413 (2015).

[110] Rao, F. *et al.* Reducing the stochasticity of crystal nucleation to enable subnanosecond memory writing. *Science* **358**, 1423–1427 (2017).

[111] Zewdie, G. M. *et al.* Chemical Design Principles for Cache-Type Sc–Sb–Te Phase-Change Memory Materials. *Chem. Mater.* **31**, 4008–4015 (2019).

[112] Matsunaga, T., Umetani, Y. & Yamada, N. Structural study of a $Ag_{3.4}In_{3.7}Sb_{76.4}Te_{16.5}$ quadruple compound utilized for phase-change optical disks. *Phys. Rev. B* **64**, 184116 (2001).

[113] Zhang, W. *et al.* How fragility makes phase-change data storage robust: Insights from ab initio simulations. *Sci. Rep.* **4**, 6529 (2014).

[114] Xu, Y. *et al.* Unraveling crystallization mechanisms and electronic structure of phase-change materials by large-scale ab initio simulations. *Adv. Mater.* **34**, 2109139 (2022).

[115] Skelton, J. M., Loke, D., Lee, T. & Elliott, S. R. Ab Initio Molecular-Dynamics Simulation of Neuromorphic Computing in Phase-Change Memory Materials. *ACS Appl. Mater. Interfaces* **7**, 14223–30 (2015).

[116] Bruns, G. *et al.* Nanosecond switching in GeTe phase change memory cells. *Appl. Phys. Lett.* **95**, 043108 (2009).

[117] Zipoli, F. & Curioni, A. Reactive potential for the study of phase-change materials: GeTe. *New J. Phys.* **15**, 123006 (2013).

[118] Deringer, V. L., Caro, M. A. & Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **31**, 1902765 (2019).

[119] Sosso, G. C., Miceli, G., Caravati, S., Behler, J. & Bernasconi, M. Neural network interatomic potential for the phase change material GeTe. *Phys. Rev. B* **85**, 174103 (2012).

[120] Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106 (2011).

[121] Sosso, G. *et al.* Fast Crystallization of the Phase Change Compound GeTe by Large-Scale Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **4**, 4241–4246 (2013).

[122] Sosso, G. C., Salvalaglio, M., Behler, J., Bernasconi, M. & Parrinello, M. Heterogeneous Crystallization of the Phase Change Material GeTe via Atomistic Simulations. *J. Phys. Chem. C* **119**, 6428–6434 (2015).

[123] Gabardi, S., Sosso, G. G., Behler, J. & Bernasconi, M. Priming effects in the crystallization of the phase change compound GeTe from atomistic simulations. *Faraday Discuss.* **213**, 287–301 (2019).

[124] Sosso, G. C., Donadio, D., Caravati, S., Behler, J. & Bernasconi, M. Thermal transport in phase-change materials from atomistic simulations. *Phys. Rev. B* **86**, 104301 (2012).

[125] Bosoni, E. *et al.* Atomistic simulations of thermal conductivity in GeTe nanowires. *J. Phys. D: Appl. Phys.* **53**, 054001 (2020).

[126] Sosso, G. C. & Bernasconi, M. Harnessing machine learning potentials to understand the functional properties of phase-change materials. *MRS Bull.* **44**, 705–709 (2019).

[127] Gabardi, S. *et al.* Atomistic Simulations of the Crystallization and Aging of GeTe Nanowires. *J. Phys. Chem. C* **121**, 23827–23838 (2017).

[128] Mitra, M., Jung, Y., Gianola, D. S. & Agarwal, R. Extremely low drift of resistance and threshold voltage in amorphous phase change nanowire devices. *Appl. Phys. Lett.* **96**, 222111 (2010).

[129] Acharya, D., Abou El Kheir, O., Campi, D. & Bernasconi, M. Crystallization kinetics of nanoconfined GeTe slabs in GeTe/$TiTe_2$-like superlattices for phase change memories. *Sci. Rep.* **14**, 3224 (2024).

[130] Fujita, T. *et al.* Pressure-induced reversal of Peierls-like distortions elicits the polyamorphic transition in GeTe and GeSe. *Nat. Commun.* **14**, 7851 (2023).

[131] Chen, Y., Campi, D., Bernasconi, M. & Mazzarello, R. Atomistic Study of the Configurational Entropy and the Fragility of Supercooled Liquid GeTe. *Adv. Funct. Mater.* **34**, 2314264 (2024).

[132] Dragoni, D., Behler, J. & Bernasconi, M. Mechanism of amorphous phase stabilization in ultrathin films of monoatomic phase change material. *Nanoscale* **13**, 16146–16155 (2021).

[133] Ronneberger, I., Chen, Y., Zhang, W. & Mazzarello, R. Local Structural Origin of the Crystallization Tendency of Pure and Alloyed Sb. *Phys. Status Solidi - Rapid Res. Lett.* **13**, 1800552 (2018).

[134] Baratella, D., Kheir, O. A. E. & Bernasconi, M. Unveiling the crystallization kinetics in Ge-rich Ge_xTe alloys by large scale simulations with a machine-learned interatomic potential (2024). `2404.15128`.

[135] Mocanu, F. C. *et al.* Modeling the Phase-Change Memory Material, $Ge_2Sb_2Te_5$, with a Machine-Learned Interatomic Potential. *J. Phys. Chem. B* **122**, 8998–9006 (2018).

[136] Mocanu, F. C., Konstantinou, K. & Elliott, S. R. Quench-rate and size-dependent behaviour in glassy $Ge_2Sb_2Te_5$ models simulated with a machine-learned Gaussian approximation potential. *J. Phys. D: Appl. Phys.* **53**, 244002 (2020).

[137] Konstantinou, K., Mocanu, F. C., Lee, T.-H. & Elliott, S. R. Revealing the intrinsic nature of the mid-gap defects in amorphous $Ge_2Sb_2Te_5$. *Nat. Commun.* **10**, 3065 (2019).

[138] Zhang, W. & Ma, E. Unveiling the structural origin to control resistance drift in phase-change memory materials. *Mater. Today* **41**, 156–176 (2020).

[139] Wang, B.-Q. *et al.* Partial melting nature of phase-change memory Ge-Sb-Te superlattice uncovered by large-scale machine learning interatomic potential molecular dynamics. *Acta Mater.* **276**, 120123 (2024).

[140] Abou El Kheir, O., Bonati, L., Parrinello, M. & Bernasconi, M. Unraveling the crystallization kinetics of the $Ge_2Sb_2Te_5$ phase change compound with a machine-learned interatomic potential. *npj Comput. Mater.* **10**, 1–12 (2024).

[141] Ko, T. W. & Ong, S. P. Recent advances and outstanding challenges for machine learning interatomic potentials. *Nat. Comput. Sci.* **3**, 998–1000 (2023).

[142] Wang, G., Sun, Y., Zhou, J. & Sun, Z. PotentialMind: Graph Convolutional Machine Learning Potential for Sb–Te Binary Compounds of Multiple Stoichiometries. *J. Phys. Chem. C* **127**, 24724–24733 (2023).

[143] Kovács, D. P. *et al.* MACE-OFF23: Transferable Machine Learning Force Fields for Organic Molecules (2023). `2312.15211`.

[144] Batatia, I. *et al.* A foundation model for atomistic materials chemistry (2024). 2401.00096.

[145] Salinga, M. *et al.* Monatomic phase change memory. *Nat. Mater.* **17**, 681–685 (2018).

[146] Konstantinou, K., Mavračić, J., Mocanu, F. C. & Elliott, S. R. Simulation of Phase-Change-Memory and Thermoelectric Materials using Machine-Learned Interatomic Potentials: Sb_2Te_3. *Phys. Status Solidi B* **258**, 2000416 (2021).

[147] Sun, L. *et al.* Ab initio molecular dynamics and materials design for embedded phase-change memory. *npj Comput. Mater.* **7**, 1–8 (2021).

[148] Cai, Z. *et al.* Physics-Based Modeling Strategies of Phase-Change Random Access Memory. *IEEE Trans. Electron Devices* **69**, 6510–6522 (2022).

[149] Smidstrup, S. *et al.* QuantumATK: An integrated platform of electronic and atomic-scale modelling tools. *J. Phys.: Condens. Matter* **32**, 015901 (2019).

[150] Elstner, M. *et al.* Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **58**, 7260–7268 (1998).

[151] Stokbro, K. *et al.* Semiempirical model for nanoscale device simulations. *Phys. Rev. B* **82**, 075420 (2010).

[152] Thesberg, M. *et al.* Monolithic TCAD simulation of phase-change memory (PCM/PRAM) + Ovonic Threshold Switch (OTS) selector device. *Solid-State Electron.* **199**, 108504 (2023).

[153] Lee, H. & Kwon, Y. Analysis of Stuck Reset Failure in Phase-Change Memory by Calculating Phase-Change Stress using Finite Element Simulation. *Phys. Status Solidi Rapid Res. Lett.* **15**, 2000419 (2021).

[154] Khan, A. I. *et al.* Ultralow-switching current density multilevel phase-change memory on a flexible substrate. *Science* **373**, 1243–1247 (2021).

[155] Simola, R. *et al.* Thermal Disturbance 3D TCAD Simulation Study in Phase Change Memory Device. In *2022 28th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, 1–4 (2022).

[156] Woods, Z., Scoggin, J., Cywar, A., Adnane, L. & Gokirmak, A. Modeling of Phase-Change Memory: Nucleation, Growth, and Amorphization Dynamics During Set and Reset: Part II—Discrete Grains. *IEEE Transactions on Electron Devices* **64**, 4472–4478 (2017).

[157] Youm, M. S., Kim, Y. T. & Sung, M. Y. Influences of partial melting and overheating on amorphization of $Ge_2Sb_2Te_5$ during the reset process. *Phys. Status Solidi A* **205**, 2657–2661 (2008).

[158] Li, Z. *et al.* Electrothermal Modeling and Design Strategies for Multibit Phase-Change Memory. *IEEE Trans. Electron Devices* **59**, 3561–3567 (2012).

[159] Foiles, S. M., Baskes, M. I. & Daw, M. S. Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. *Phys. Rev. B* **33**, 7983–7991 (1986).

[160] Daw, M. S. & Baskes, M. I. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Phys. Rev. B* **29**, 6443–6453 (1984).

[161] Friederich, P., Häse, F., Proppe, J. & Aspuru-Guzik, A. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.* **20**, 750–761 (2021).

[162] Behler, J. & Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).

[163] Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. & Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).

[164] Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).

[165] Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).

[166] Zhang, L., Han, J., Wang, H., Car, R. & E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).

[167] Artrith, N. & Kolpak, A. M. Understanding the composition and activity of electrocatalytic nanoalloys in aqueous solvents: A combination of DFT and accurate neural network potentials. *Nano Lett.* **14**, 2670–2676 (2014).

[168] Wang, C., Aoyagi, K., Wisesa, P. & Mueller, T. Lithium Ion Conduction in Cathode Coating Materials from On-the-Fly Machine Learning. *Chem. Mater.* **32**, 3741–3752 (2020).

[169] Musil, F. *et al.* Physics-Inspired Structural Representations for Molecules and Materials. *Chem. Rev.* **121**, 9759–9815 (2021).

[170] Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).

[171] Deringer, V. L. *et al.* Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **121**, 10073–10141 (2021).

[172] Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).

[173] Erhard, L. C., Rohrer, J., Albe, K. & Deringer, V. L. Modelling atomic and nanoscale structure in the silicon–oxygen system through active machine learning. *Nat. Commun.* **15**, 1927 (2024).

[174] Bochkarev, A. *et al.* Efficient parametrization of the atomic cluster expansion. *Phys. Rev. Mater.* **6**, 013804 (2022).

[175] Batzner, S. *et al.* E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).

[176] Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **95**, 094203 (2017).

[177] Bartók, A. P., Kermode, J., Bernstein, N. & Csányi, G. Machine Learning a General-Purpose Interatomic Potential for Silicon. *Phys. Rev. X* **8**, 041048 (2018).

[178] Deringer, V. L., Caro, M. A. & Csányi, G. A general-purpose machine-learning force field for bulk and nanostructured phosphorus. *Nat. Commun.* **11**, 5461 (2020).

[179] Erhard, L. C., Rohrer, J., Albe, K. & Deringer, V. L. A machine-learned interatomic potential for silica and its relation to empirical models. *npj Comput. Mater.* **8**, 90 (2022).

[180] Zhou, Y., Zhang, W., Ma, E. & Deringer, V. L. Device-scale atomistic modelling of phase-change memory materials. *Nat. Electron.* **6**, 746–754 (2023).

[181] Solak, E., Murray-smith, R., Leithead, W., Leith, D. & Rasmussen, C. Derivative observations in gaussian process models of dynamic systems. In Becker, S., Thrun, S. & Obermayer, K. (eds.) *Advances in Neural Information Processing Systems*, vol. 15 (MIT Press, 2002).

[182] Lysogorskiy, Y. *et al.* Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. *npj Comput. Mater.* **7**, 1–12 (2021).

[183] Qamar, M., Mrovec, M., Lysogorskiy, Y., Bochkarev, A. & Drautz, R. Atomic Cluster Expansion for Quantum-Accurate Large-Scale Simulations of Carbon. *J. Chem. Theory Comput.* **19**, 5151–5167 (2023).

[184] Liang, Y. *et al.* Atomic cluster expansion for Pt–Rh catalysts: From ab initio to the simulation of nanoclusters in few steps. *J. Mater. Res.* **38**, 5125–5135 (2023).

[185] Ibrahim, E., Lysogorskiy, Y., Mrovec, M. & Drautz, R. Atomic cluster expansion for a general-purpose interatomic potential of magnesium. *Phys. Rev. Mater.* **7**, 113801 (2023).

[186] Finnis, M. W. & Sinclair, J. E. A simple empirical N-body potential for transition metals. *Philos. Mag. A* **50**, 45–55 (1984).

[187] Ben Mahmoud, C., Gardner, J. L. A. & Deringer, V. L. Data as the next challenge in atomistic machine learning. *Nat. Comput. Sci.* **4**, 384–387 (2024).

[188] Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **56**, 12828–12840 (2017).

[189] Unke, O. T. *et al.* Machine Learning Force Fields. *Chem. Rev.* **121**, 10142–10186 (2021).

[190] Blum, L. C. & Reymond, J.-L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *J. Am. Chem. Soc.* **131**, 8732–8733 (2009).

[191] Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **108**, 058301 (2012).

[192] Noé, F., Olsson, S., Köhler, J. & Wu, H. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science* **365**, eaaw1147 (2019).

[193] Cheng, B., Mazzola, G., Pickard, C. J. & Ceriotti, M. Evidence for supercritical behaviour of high-pressure liquid hydrogen. *Nature* **585**, 217–220 (2020).

[194] Szlachta, W. J., Bartók, A. P. & Csányi, G. Accuracy and transferability of Gaussian approximation potential models for tungsten. *Phys. Rev. B* **90**, 104108 (2014).

[195] Caro, M. A., Deringer, V. L., Koskinen, J., Laurila, T. & Csányi, G. Growth Mechanism and Origin of High sp^3 Content in Tetrahedral Amorphous Carbon. *Phys. Rev. Lett.* **120**, 166101 (2018).

[196] Caro, M. A., Csányi, G., Laurila, T. & Deringer, V. L. Machine learning driven simulated deposition of carbon films: From low-density to diamondlike amorphous carbon. *Phys. Rev. B* **102**, 174201 (2020).

[197] Bernstein, N., Csányi, G. & Deringer, V. L. De novo exploration and self-guided learning of potential-energy surfaces. *npj Comput. Mater.* **5**, 1–9 (2019).

[198] Pickard, C. J. & Needs, R. J. High-Pressure Phases of Silane. *Phys. Rev. Lett.* **97**, 045504 (2006).

[199] Pickard, C. J. & Needs, R. J. Ab initio random structure searching. *J. Phys.: Condens. Matter* **23**, 053201 (2011).

[200] Deringer, V. L., Csányi, G. & Proserpio, D. M. Extracting Crystal Chemistry from Amorphous Carbon Structures. *ChemPhysChem* **18**, 873–877 (2017).

[201] Deringer, V. L., Proserpio, D. M., Csányi, G. & Pickard, C. J. Data-driven learning and prediction of inorganic crystal structures. *Faraday Discuss.* **211**, 45–59 (2018).

[202] Deringer, V. L., Pickard, C. J. & Csányi, G. Data-driven learning of total and local energies in elemental boron. *Phys. Rev. Lett.* **120**, 156001 (2018).

[203] Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V. & Oganov, A. R. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys. Rev. B* **99**, 064114 (2019).

[204] Hajinazar, S., Shao, J. & Kolmogorov, A. N. Stratified construction of neural network based interatomic models for multicomponent materials. *Phys. Rev. B* **95**, 014114 (2017).

[205] Mahoney, M. W. & Drineas, P. CUR matrix decompositions for improved data analysis. *Proc. Natl. Acad. Sci.* **106**, 697–702 (2009).

[206] Mavračić, J., Mocanu, F. C., Deringer, V. L., Csányi, G. & Elliott, S. R. Similarity Between Amorphous and Crystalline Phases: The Case of TiO_2. *J. Phys. Chem. Lett.* **9**, 2985–2990 (2018).

[207] Thomas du Toit, D. F., Zhou, Y. & Deringer, V. L. Hyperparameter Optimization for Atomic Cluster Expansion Potentials (2024). 2408.00656.

[208] Thomas du Toit, D. F. & Deringer, V. L. Cross-platform hyperparameter optimization for machine learning interatomic potentials. *J. Chem. Phys.* **159**, 024803 (2023).

[209] Morrow, J. D., Gardner, J. L. A. & Deringer, V. L. How to validate machine-learned interatomic potentials. *J. Chem. Phys.* **158**, 121501 (2023).

[210] Wang, J. *et al.* Enhancing thermoelectric performance of Sb_2Te_3 through swapped bilayer defects. *Nano Energy* **79**, 105484 (2021).

[211] Zhou, Y., Kirkpatrick, W. & Deringer, V. L. Cluster Fragments in Amorphous Phosphorus and their Evolution under Pressure. *Adv. Mater.* **34**, 2107515 (2022).

[212] Zhou, Y., Elliott, S. R. & Deringer, V. L. Structure and Bonding in Amorphous Red Phosphorus. *Angew. Chem. Int. Ed.* **62**, e202216658 (2023).

[213] Kim, Y. *et al.* An amorphous red phosphorus/carbon composite as a promising anode material for sodium ion batteries. *Adv. Mater.* **25**, 3045–3049 (2013).

[214] Wu, X. *et al.* Enhanced capacity of chemically bonded phosphorus/carbon composite as an anode material for potassium-ion batteries. *J. Power Sources* **378**, 460–467 (2018).

[215] Jiao, X. *et al.* Amorphous phosphorus-carbon nanotube hybrid anode with ultralong cycle life and high-rate capability for lithium-ion batteries. *Carbon* **148**, 518–524 (2019).

[216] Liu, S. *et al.* Nanoporous Red Phosphorus on Reduced Graphene Oxide as Superior Anode for Sodium-Ion Batteries. *ACS Nano* **12**, 7380–7387 (2018).

[217] Capone, I., Hurlbutt, K., Naylor, A. J., Xiao, A. W. & Pasta, M. Effect of the Particle-Size Distribution on the Electrochemical Performance of a Red Phosphorus-Carbon Composite Anode for Sodium-Ion Batteries. *Energy Fuels* **33**, 4651–4658 (2019).

[218] Liu, Y. *et al.* Red-phosphorus-impregnated carbon nanofibers for sodium-ion batteries and liquefaction of red phosphorus. *Nat. Commun.* **11**, 2520 (2020).

[219] Zhao, S., Kang, W. & Xue, J. The potential application of phosphorene as an anode material in Li-ion batteries. *J. Mater. Chem. A* **2**, 19046–19052 (2014).

[220] Mayo, M., Griffith, K. J., Pickard, C. J. & Morris, A. J. Ab Initio Study of Phosphorus Anodes for Lithium- and Sodium-Ion Batteries. *Chem. Mater.* **28**, 2011–2021 (2016).

[221] Lannin, J. S., Shanabrook, B. V. & Gompf, F. Intermediate range order in amorphous phosphorous. *J. Non-Cryst. Solids* **49**, 209–219 (1982).

[222] Lannin, J. S. & Shanabrook, B. V. Raman scattering and infrared absorption in bulk amorphous red phosphorous. *Solid State Commun.* **28**, 497–500 (1978).

[223] Yang, Z. *et al.* Field-effect transistors based on amorphous black phosphorus ultrathin films by pulsed laser deposition. *Adv. Mater.* **27**, 3748–54 (2015).

[224] Elliott, S. R., Dore, J. C. & Marseglia, E. The Structure of Amorphous Phosphorus. *J. phys., Colloq.* **46**, C8 (1985).

[225] Thurn, H. & Krebs, H. Crystal Structure of Violet Phosphorus. *Angew. Chem. Int. Ed.* **5**, 1047–1048 (1966).

[226] Ruck, M. *et al.* Fibrous red phosphorus. *Angew. Chem. Int. Ed.* **44**, 7616–7619 (2005).

[227] Olego, D. J. *et al.* The microscopic structure of bulk amorphous red phosphorus: A Raman scattering investigation. *Solid State Commun.* **52**, 311–314 (1984).

[228] Fasol, G., Cardona, M., Hönle, W. & von Schnering, H. G. Lattice dynamics of Hittorf's phosphorus and identification of structural groups and defects in amorphous red phosphorus. *Solid State Commun.* **52**, 307–310 (1984).

[229] Zaug, J. M., Soper, A. K. & Clark, S. M. Pressure-dependent structures of amorphous red phosphorus and the origin of the first sharp diffraction peaks. *Nat. Mater.* **7**, 890–899 (2008).

[230] Jones, R. O. & Hohl, D. Structure of phosphorus clusters using simulated annealing—P2 to P8. *J. Chem. Phys.* **92**, 6710–6721 (1990).

[231] Hohl, D. & Jones, R. O. Amorphous phosphorus: A cluster-network model. *Phys. Rev. B* **45**, 8995–9005 (1992).

[232] Böcker, S. & Häser, M. Covalent Structures of Phosphorus: A Comprehensive Theoretical Study. *Z. Anorg. Allg. Chem.* **621**, 258–286 (1995).

[233] Baudler, M. Chain and Ring Phosphorus Compounds—Analogies between Phosphorus and Carbon Chemistry. *Angew. Chem. Int. Ed.* **21**, 492–512 (1982).

[234] von Schnering, H. G. Homoatomic Bonding of Main Group Elements. *Angew. Chem. Int. Ed.* **20**, 33–51 (1981).

[235] Impellizzeri, A. *et al.* Simulated Raman spectra of bulk and low-dimensional phosphorus allotropes. *Phys. Chem. Chem. Phys.* **23**, 16611–16622 (2021).

[236] Deringer, V. L. *et al.* Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics. *J. Phys. Chem. Lett.* **9**, 2879–2885 (2018).

[237] Katayama, Y. *et al.* A first-order liquid-liquid phase transition in phosphorus. *Nature* **403**, 170–173 (2000).

[238] Zeng, Z. *et al.* Synthesis of quenchable amorphous diamond. *Nat. Commun.* **8**, 322 (2017).

[239] de Tomas, C. *et al.* Structural prediction of graphitization and porosity in carbide-derived carbons. *Carbon* **119**, 1–9 (2017).

[240] Deringer, V. L. *et al.* Towards an atomistic understanding of disordered carbon electrode materials. *Chem. Commun.* **54**, 5988–5991 (2018).

[241] McKenzie, D. R., Muller, D. & Pailthorpe, B. A. Compressive-stress-induced formation of thin-film tetrahedral amorphous carbon. *Phys. Rev. Lett.* **67**, 773–776 (1991).

[242] Deb, S. K., Wilding, M., Somayazulu, M. & McMillan, P. F. Pressure-induced amorphization and an amorphous-amorphous transition in densified porous silicon. *Nature* **414**, 528–530 (2001).

[243] Le Roux, S. & Jund, P. Ring statistics analysis of topological networks: New approach and application to amorphous GeS_2 and SiO_2 systems. *Comput. Mater. Sci.* **49**, 70–83 (2010).

[244] Xiang, H. *et al.* The mechanism of structural changes and crystallization kinetics of amorphous red phosphorus to black phosphorus under high pressure. *Chem. Commun.* **55**, 8094–8097 (2019).

[245] Rissi, E. N., Soignard, E., McKiernan, K. A., Benmore, C. J. & Yarger, J. L. Pressure-induced crystallization of amorphous red phosphorus. *Solid State Commun.* **152**, 390–394 (2012).

[246] Brazhkin, V. V. & Zerr, A. J. Relative stability of red and black phosphorus at P<1 GPa. *J. Mater. Sci.* **27**, 2677–2681 (1992).

[247] Brown, A. & Rundqvist, S. Refinement of the crystal structure of black phosphorus. *Acta Crystallogr.* **19**, 684–685 (1965).

[248] Pfitzner, A., Bräu, M. F., Zweck, J., Brunklaus, G. & Eckert, H. Phosphorus nanorods–two allotropic modifications of a long-known element. *Angew. Chem. Int. Ed.* **43**, 4228–31 (2004).

[249] Bachhuber, F. *et al.* The extended stability range of phosphorus allotropes. *Angew. Chem. Int. Ed.* **53**, 11629–11633 (2014).

[250] Falconi, S., Crichton, W. A., Mezouar, M., Monaco, G. & Nardone, M. Structure of fluid phosphorus at high temperature and pressure: An x-ray diffraction study. *Phys. Rev. B* **70**, 144109 (2004).

[251] Monaco, G., Falconi, S., Crichton, W. A. & Mezouar, M. Nature of the first-order phase transition in fluid phosphorus at high temperature and pressure. *Phys. Rev. Lett.* **90**, 255701 (2003).

[252] Scelta, D. *et al.* Interlayer Bond Formation in Black Phosphorus at High Pressure. *Angew. Chem. Int. Ed.* **56**, 14135–14140 (2017).

[253] Farman, H., Blakey, D., Dore, J. C., Bellissent-Funel, M. C. & Elliott, S. R. Neutron diffraction studies of network structure in amorphous phosphorus. *Phys. Scr.* **49**, 634–636 (1994).

[254] Elliott, S. R. The origin of the first sharp diffraction peak in the structure factor of covalent glasses and liquids. *J. Phys.: Condens. Matter* **4**, 7661–7678 (1992).

[255] Elliott, S. R. Medium-range structural order in covalent amorphous solids. *Nature* **354**, 445–452 (1991).

[256] Farman, H., Dore, J. C. & Elliott, S. R. Intermediate-range order in amorphous phosphorus. *Phys. Lett. A* **186**, 410–414 (1994).

[257] Stukowski, A. Computational Analysis Methods in Atomistic Modeling of Crystals. *JOM* **66**, 399–407 (2013).

[258] Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).

[259] Katayama, Y. *et al.* Macroscopic separation of dense fluid phase and liquid phase of phosphorus. *Science* **306**, 848–851 (2004).

[260] Deringer, V. L., Pickard, C. J. & Proserpio, D. M. Hierarchically Structured Allotropes of Phosphorus from Data-Driven Exploration. *Angew. Chem. Int. Ed.* **59**, 15880–15885 (2020).

[261] Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).

[262] Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).

[263] Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).

[264] Aykol, M., Doak, J. W. & Wolverton, C. Phosphorus allotropes: Stability of black versus red phosphorus re-examined by means of the van der Waals inclusive density functional method. *Phys. Rev. B* **95**, 214115 (2017).

[265] Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

[266] Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).

[267] Tkatchenko, A., DiStasio, R. A., Car, R. & Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **108**, 236402 (2012).

[268] Ambrosetti, A., Reilly, A. M., DiStasio, R. A. & Tkatchenko, A. Long-range correlation energy calculated from coupled atomic response functions. *J. Chem. Phys.* **140**, 18A508 (2014).

[269] Sansone, G. *et al.* On the exfoliation and anisotropic thermal expansion of black phosphorus. *Chem. Commun.* **54**, 9793–9796 (2018).

[270] Shulenburger, L., Baczewski, A. D., Zhu, Z., Guan, J. & Tománek, D. The Nature of the Interlayer Interaction in Bulk and Few-Layer Phosphorus. *Nano Lett.* **15**, 8170–8175 (2015).

[271] Schutz, M., Maschio, L., Karttunen, A. J. & Usvyat, D. Exfoliation Energy of Black Phosphorus Revisited: A Coupled Cluster Benchmark. *J. Phys. Chem. Lett.* **8**, 1290–1294 (2017).

[272] Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).

[273] Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

[274] Simon, A., Borrmann, H. & Horakh, J. On the Polymorphism of White Phosphorus. *Chem. Ber.* **130**, 1235–1240 (2006).

[275] Bachhuber, F. *et al.* Van der Waals interactions in selected allotropes of phosphorus. *Z. Kristallogr. Cryst. Mater.* **230**, 107–115 (2015).

[276] Medvedev, N. N., Voloshin, V. P., Luchnikov, V. A. & Gavrilova, M. L. An algorithm for three-dimensional Voronoi S-network. *J. Comput. Chem.* **27**, 1676–1692 (2006).

[277] De, S., Bartók, A. P., Csányi, G. & Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Phys. Chem. Chem. Phys.* **18**, 13754–13769 (2016).

[278] Zhou, Y. *et al.* Bonding similarities and differences between Y–Sb–Te and Sc–Sb–Te phase-change memory materials. *J. Mater. Chem. C* **8**, 3646–3654 (2020).

[279] Deringer, V. L. *et al.* Origins of structural and electronic transitions in disordered silicon. *Nature* **589**, 59–64 (2021).

[280] Oh, H. S. *et al.* Element-resolved local lattice distortion in complex concentrated alloys: An observable signature of electronic effects. *Acta Mater.* **216**, 117135 (2021).

[281] Orava, J., Greer, A. L., Gholipour, B., Hewak, D. W. & Smith, C. E. Characterization of supercooled liquid $Ge_2Sb_2Te_5$ and its crystallization by ultrafast-heating calorimetry. *Nat. Mater.* **11**, 279–83 (2012).

[282] Zhou, Y., Elliott, S. R., du Toit, D. F. T., Zhang, W. & Deringer, V. L. The pathway to chirality in elemental tellurium (2024). `2409.03860`.

[283] Dong, Z. & Ma, Y. Atomic-level handedness determination of chiral crystals using aberration-corrected scanning transmission electron microscopy. *Nat. Commun.* **11**, 1588 (2020).

[284] Ben-Moshe, A. *et al.* Enantioselective control of lattice and shape chirality in inorganic nanostructures using chiral biomolecules. *Nat. Commun.* **5**, 4302 (2014).

[285] Shtukenberg, A. G., Punin, Y. O., Gujral, A. & Kahr, B. Growth actuated bending and twisting of single crystals. *Angew. Chem. Int. Ed.* **53**, 672–699 (2014).

[286] di Gregorio, M. C. *et al.* Emergence of chirality and structural complexity in single crystals at the molecular and morphological levels. *Nat. Commun.* **11**, 380 (2020).

[287] Wang, P.-p., Yu, S.-J., Govorov, A. O. & Ouyang, M. Cooperative expression of atomic chirality in inorganic nanostructures. *Nat. Commun.* **8**, 14312 (2017).

[288] Monego, D. *et al.* Ligand-induced incompatible curvatures control ultrathin nanoplatelet polymorphism and chirality. *Proc. Natl. Acad. Sci.* **121**, e2316299121 (2024).

[289] Gonzalez, A. V., Gonzalez, M. & Hanrath, T. Emergence and inversion of chirality in hierarchical assemblies of CdS nanocrystal fibers. *Sci. Adv.* **9**, eadi5520 (2023).

[290] Im, S. W. *et al.* Investigating chiral morphogenesis of gold using generative cellular automata. *Nat. Mater.* **23**, 977–983 (2024).

[291] Ben-Moshe, A. *et al.* The chain of chirality transfer in tellurium nanocrystals. *Science* **372**, 729–733 (2021).

[292] Gatti, G. *et al.* Radial spin texture of the Weyl fermions in chiral tellurium. *Phys. Rev. Lett.* **125**, 216402 (2020).

[293] Calavalle, F. *et al.* Gate-tuneable and chirality-dependent charge-to-spin conversion in tellurium nanowires. *Nat. Mater.* **21**, 526–532 (2022).

[294] Jnawali, G. *et al.* Ultrafast photoinduced band splitting and carrier dynamics in chiral tellurium nanosheets. *Nat. Commun.* **11**, 3991 (2020).

[295] Li, L. *et al.* Dual sensing signal decoupling based on tellurium anisotropy for VR interaction and neuro-reflex system application. *Nat. Commun.* **13**, 5975 (2022).

[296] Yang, Y. *et al.* A new opportunity for the emerging tellurium semiconductor: Making resistive switching devices. *Nat. Commun.* **12**, 6081 (2021).

[297] He, Z., Yang, Y., Liu, J.-W. & Yu, S.-H. Emerging tellurium nanostructures: Controllable synthesis and their applications. *Chem. Soc. Rev.* **46**, 2732–2753 (2017).

[298] Li, S. *et al.* Programmable nucleation and growth of ultrathin tellurium nanowires via a pulsed physical vapor deposition design. *Adv. Funct. Mater.* **33**, 2211527 (2023).

[299] Zhao, C. *et al.* Tellurium single-crystal arrays by low-temperature evaporation and crystallization. *Adv. Mater.* **33**, e2100860 (2021).

[300] Keller, H. & Stuke, J. Elektrische und optische Eigenschaften von amorphem Tellur. *Phys. Status Solidi B* **8**, 831–840 (1965).

[301] Shen, J. *et al.* Elemental electrical switch enabling phase segregation-free operation. *Science* **374**, 1390–1394 (2021).

[302] Epstein, A. S., Fritzsche, H. & Lark-Horovitz, K. Electrical properties of tellurium at the melting point and in the liquid state. *Phys. Rev.* **107**, 412–419 (1957).

[303] Ma, Y., Oleynikov, P. & Terasaki, O. Electron crystallography for determining the handedness of a chiral zeolite nanocrystal. *Nat. Mater.* **16**, 755–759 (2017).

[304] Decker, A., Landrum, G. A. & Dronskowski, R. Structural and electronic Peierls distortions in the elements (A): The crystal structure of tellurium. *Z. Anorg. Allg. Chem.* **628**, 295–302 (2002).

[305] Burdett, J. K. & McLarnan, T. J. A study of the arsenic, black phosphorus, and other structures derived from rock salt by bond-breaking processes. I. Structural enumeration. *J. Chem. Phys.* **75**, 5764–5773 (1981).

[306] Champness, C. H. & Hoffmann, R. H. Conductivity changes associated with the crystallization of amorphous selenium. *J. Non-Cryst. Solids* **4**, 138–148 (1970).

[307] Šimečková, M. & Hrubý, A. A study on the crystallization of amorphous arsenic. *Mater. Res. Bull.* **12**, 65–72 (1977).

[308] Knights, J. C. & Mahan, J. E. Optical and electrical properties of amorphous arsenic. *Solid State Commun.* **21**, 983–986 (1977).

[309] Jamieson, J. C. Crystal structures adopted by black phosphorus at high pressures. *Science* **139**, 1291–1292 (1963).

[310] Beister, H. J., Strössner, K. & Syassen, K. Rhombohedral to simple-cubic phase transition in arsenic under pressure. *Phys. Rev. B* **41**, 5535–5543 (1990).

[311] Akahama, Y., Kobayashi, M. & Kawamura, H. Structural studies of pressure-induced phase transitions in selenium up to 150 GPa. *Phys. Rev. B* **47**, 20–26 (1993).

[312] Hunderi, O. Optical properties of crystalline and amorphous bismuth films. *J. Phys. F: Met. Phys.* **5**, 2214 (1975).

[313] Hegedus, J. & Elliott, S. R. Computer-simulation design of new phase-change memory materials. *Phys. Status Solidi A* **207**, 510–515 (2010).

[314] Shen, X. *et al.* Surface effects on the crystallization kinetics of amorphous antimony. *Nanoscale* **15**, 15259–15267 (2023).

[315] Da Silva, J., Walsh, A. & Lee, H. Insights into the structure of the stable and metastable $(GeTe)_m(Sb_2Te_3)_n$ compounds. *Phys. Rev. B* **78**, 224111 (2008).

[316] Matsunaga, T., Kubota, Y. & Yamada, N. Structures of stable and metastable $Ge_2Sb_2Te_5$, an intermetallic compound in the GeTe-Sb_2Te_3 pseudobinary systems. *Acta Crystallogr., Sect. B* **60**, 685 (2004).

[317] Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **91**, 146401 (2003).

[318] Akola, J., Jones, R. O., Kohara, S., Usuki, T. & Bychkov, E. Density variations in liquid tellurium: Roles of rings, chains, and cavities. *Phys. Rev. B* **81**, 094202 (2010).

[319] Akola, J. & Jones, R. O. Structure and dynamics in amorphous tellurium and Te_n clusters: A density functional study. *Phys. Rev. B* **85**, 134103 (2012).

[320] Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).

[321] Sun, P. *et al.* Structural changes across thermodynamic maxima in supercooled liquid tellurium: A water-like scenario. *Proc. Natl. Acad. Sci.* **119**, e2202044119 (2022).

[322] Thompson, A. P. *et al.* LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **271**, 108171 (2022).

[323] Ichikawa, T. Electron diffraction study of the local atomic arrangement in amorphous tellurium films. *Phys. Status Solidi B* **56**, 707–715 (1973).

[324] Sun, Y. *et al.* Nanosecond phase-transition dynamics in elemental tellurium. *Adv. Funct. Mater.* 2408725 (2024).

[325] Kühne, T. D., Krack, M., Mohamed, F. R. & Parrinello, M. Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.* **98**, 066401 (2007).

[326] Kühne, T. D. *et al.* CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **152**, 194103 (2020).

[327] Goedecker, S., Teter, M. & Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703 (1996).

[328] Abbas, H. G. & Hahn, J. R. Crystallization mechanism of liquid tellurium from classical molecular dynamics simulation. *Mater. Chem. Phys.* **240**, 122235 (2020).

[329] Vladimirovich, K. A. & Nikolaevna, G. O. Chirality properties of modeling water in different aqueous systems. In *Chirality from Molecular Electronic States* (IntechOpen, 2018).

[330] Bonald, T., Charpentier, B., Galland, A. & Hollocou, A. Hierarchical graph clustering using node pair sampling (2018). `1806.01664`.

[331] El-Machachi, Z., Wilson, M. & Deringer, V. L. Exploring the configurational space of amorphous graphene with machine-learned atomic energies. *Chem. Sci.* **13**, 13720–13731 (2022).

[332] Morrow, J. D. *et al.* Understanding defects in amorphous silicon with million-atom simulations and machine learning. *Angew. Chem. Int. Ed.* **136**, e202403842 (2024).

[333] Zhu, J. *et al.* Formation of chiral branched nanowires by the Eshelby Twist. *Nat. Nanotechnol.* **3**, 477–481 (2008).

[334] Liu, Y. *et al.* Helical van der Waals crystals with discretized Eshelby twist. *Nature* **570**, 358–362 (2019).

[335] Morin, S. A., Bierman, M. J., Tong, J. & Jin, S. Mechanism and kinetics of spontaneous nanotube growth driven by screw dislocations. *Science* **328**, 476–480 (2010).

[336] Bierman, M. J., Lau, Y. K. A., Kvit, A. V., Schmitt, A. L. & Jin, S. Dislocation-driven nanowire growth and Eshelby twist. *Science* **320**, 1060–1063 (2008).

[337] Wang, X.-D., Zhang, W. & Ma, E. Monatomic phase-change switch. *Sci. Bull.* **67**, 888–890 (2022).

[338] Zhang, L. *et al.* One-selector one-resistor cross-point array with threshold switching selector. *IEEE Trans. Electron Devices* **62**, 3250–3257 (2015).

[339] Wuttig, M. & Yamada, N. Phase-change materials for rewriteable data storage. *Nat. Mater.* **6**, 824–832 (2007).

[340] Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R. & Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020).

[341] Liu, B. *et al.* Multi-level phase-change memory with ultralow power consumption and resistance drift. *Sci. Bull.* **66**, 2217–2224 (2021).

[342] Shastri, B. J. *et al.* Photonics for artificial intelligence and neuromorphic computing. *Nat. Photon.* **15**, 102–114 (2021).

[343] Akola, J. & Jones, R. Structural phase transitions on the nanoscale: The crucial pattern in the phase-change materials $Ge_2Sb_2Te_5$ and GeTe. *Phys. Rev. B* **76**, 235201 (2007).

[344] Xu, M., Cheng, Y., Sheng, H. & Ma, E. Nature of Atomic Bonding and Atomic Structure in the Phase-Change $Ge_2Sb_2Te_5$ Glass. *Phys. Rev. Lett.* **103**, 195502 (2009).

[345] Huang, B. & Robertson, J. Bonding origin of optical contrast in phase-change memory materials. *Phys. Rev. B* **81**, 081204(R) (2010).

[346] Raty, J.-Y. *et al.* A Quantum-Mechanical Map for Bonding and Properties in Solids. *Adv. Mater.* **31**, 1806280 (2019).

[347] Shi, M., Li, J., Tao, M., Zhang, X. & Liu, J. Artificial intelligence model for efficient simulation of monatomic phase change material antimony. *Mater. Sci. Semicond. Process.* **136**, 106146 (2021).

[348] Jinnouchi, R., Lahnsteiner, J., Karsai, F., Kresse, G. & Bokdam, M. Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference. *Phys. Rev. Lett.* **122**, 225701 (2019).

[349] Clark, S. J. *et al.* First principles methods using CASTEP. *Z. Kristallogr. Cryst. Mater.* **220**, 567–570 (2005).

[350] Perdew, J. *et al.* Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).

[351] Lee, D., Lee, K., Yoo, D., Jeong, W. & Han, S. Crystallization of amorphous GeTe simulated by neural network potential addressing medium-range order. *Comput. Mater. Sci.* **181**, 109725 (2020).

[352] Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

[353] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).

[354] Shinoda, W., Shiga, M. & Mikami, M. Rapid estimation of elastic constants by molecular dynamics simulation under constant stress. *Phys. Rev. B* **69**, 134103 (2004).

[355] Song, Z. *et al.* 12-state multi-level cell storage implemented in a 128 Mb phase change memory chip. *Nanoscale* **13**, 10455–10461 (2021).

[356] Tuma, T., Pantazi, A., Le Gallo, M., Sebastian, A. & Eleftheriou, E. Stochastic phase-change neurons. *Nat. Nanotechnol.* **11**, 693–699 (2016).

[357] Waldecker, L. *et al.* Time-domain separation of optical properties from structural transitions in resonantly bonded materials. *Nat. Mater.* **14**, 991–5 (2015).

[358] Cobelli, M., Galante, M., Gabardi, S., Sanvito, S. & Bernasconi, M. First-Principles Study of Electromigration in the Metallic Liquid State of GeTe and Sb_2Te_3 Phase-Change Compounds. *J. Phys. Chem. C* **124**, 9599–9603 (2020).

[359] McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).

[360] Müller, M. J. *et al.* Decoupling Nucleation and Growth in Fast Crystallization of Phase Change Materials. *Adv. Funct. Mater.* **34**, 2403476 (2024).

[361] Li, K., Liu, B., Zhou, J. & Sun, Z. Revealing the crystallization dynamics of Sb–Te phase change materials by large-scale simulations. *J. Mater. Chem. C* **12**, 3897–3906 (2024).

[362] Lian, X. & Wang, L. Boolean Logic Function Realized by Phase-Change Blade Type Random Access Memory. *IEEE Trans. Electron Devices* **69**, 1849–1857 (2022).