Phase-Change Superlattice Materials toward Low Power Consumption and High Density Data Storage: Microscopic Picture, Working Principles, and Optimization

Xian-Bin Li, Nian-Ke Chen,* Xue-Peng Wang, and Hong-Bo Sun*

To meet the requirement of big data era and neuromorphic computations, nonvolatile memory with fast speed, high density, and low power consumption is urgently needed. As an emerging technology, phase-change memory is a promising candidate to solve this problem. However, the drawback of the high power consumption hinders their applications. Most recently, a new phase-change material of [(GeTe)\textsubscript{n}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{y}] superlattice attracts intensive attentions owing to its ultralow power consumption comparing with conventional phase-change memory devices. Many studies on this new material have been reported. However, there still lacks a comprehensive and unified understanding of its atomic picture and working mechanism. This article at first summarizes the broad applications for phase-change materials. Then, the major progresses of phase-change superlattices to understand the microscopic structure and working principles for data storage are discussed. Strategies on material optimizations to further enhance device performances are proposed. Finally, an outlook on new applications with these advanced superlattice materials is suggested.

1. Introduction

Information storage has played significant roles in the evolution of human's history. Nowadays, the development of electronic technology dramatically increases the amount of digital data. It has been announced that the size of digital data in the global world doubles every two years, and by 2020, the size of digital data will reach 44 zettabytes.\[1\] Figure 1 shows the growth of digital data in recent years and the prediction. These data are mainly generated by automated machines and electronic devices, such as automobiles, computers, sensors, and smartphones. With the development of Internet of Things, an extremely large amount of data are being generated and transferred at every second in the form of video, music, picture, online social networking, business information, and so on. Thus, the storage, transmission, and processing for such big data will face serious challenges.\[2-6\] It has been proposed that big data will cost large processing resources and thus need a hardware revolution.\[7,8\] Nonvolatile memory (NVM) devices with fast speed, high density, and low power consumption are urgently needed to handle these problems.\[9,10\]

Figure 2 shows the typical hierarchy of memory and storage in computers. The Flash, hard disk drive (HDD), and compact disc (CD)/digital video disc (DVD)/blu-ray disc (BD) are nonvolatile, while static random access memory (SRAM) and dynamic random access memory (DRAM) are volatile. That means the former will retain data even without power but the latter will lose data without power. As a popular NVM, Flash has been widely used in mobile phones, universal serial bus (USB) flash drives, and other portable electronic devices. Nowadays, Flash-based solid state disk is replacing the traditional HDD because it has a much faster speed and a lower power consumption. However, there still exists a speed gap between the nonvolatile and volatile memories. The write/erase time of Flash is about 10\textsuperscript{3} ns while the write/erase time of DRAM is only 10\textsuperscript{2} ns.\[11\] Also, the endurance of Flash is only 10\textsuperscript{3} cycles (for multilevel cell, MLC) or 10\textsuperscript{5} cycles (for single-level cell), which is still too low for big data era.\[12\] Phase-change memory (PCM) has been proposed to be an emerging memory which can cover a wide-range technology from tertiary storage to off-chip memory (the community says that a universal memory).\[13,14\] It not only makes up the speed gap (with a working speed of ≈30 ns) but also possesses a high areal density (up to 200 Gb in.\textsuperscript{2}) and a better endurance (which was predicted up to 10\textsuperscript{12} cycles).\[15\] Other performances of the PCM, such as data retention, recording density, and MLC capability, are also very good. As such, phase-change RAM (PCRAM) is a promising candidate to replace Flash and will possibly become the mainstream NVM in the future.

This technology was originally proposed by Ovshinsky in the 1960s.\[11\] Its working principle relies on the reversible switching between crystalline and amorphous states (Figure 3a–d).
Figure 3a,b shows the transmission electron microscope (TEM) cross-sections of the SET and RESET states of a typical mushroom-type device, respectively. As illustrated by the schematics in the expanded green and orange squares, the SET state is a crystalline phase with a high optical reflectivity and a low electrical resistivity while the RESET state is an amorphous phase with a low optical reflectivity and a high electrical resistivity. Generally, a large-intensity but short-duration optical/electrical pulse (orange curve in Figure 3c) heats the material over its melting point and turns it into a liquid. Then, the subsequently rapid quenching freezes the liquid into an amorphous phase. By contrast, a mild-intensity but long-duration pulse (green curve in Figure 3d) heats the material over its crystallization temperature but below its melting point. Then, the amorphous phase gradually recrystallizes. In the device, the bottom electrode usually serves as a heater. The mushroom-shaped amorphous region near the bottom electrode is clearly seen in a TEM cross-section (shown with yellow semi-circle in Figure 3b). Recently, electrical fields induced motions of dislocations in nanowires and electronic excitation induced solid-to-solid transition have also been suggested as the phase-transition mechanisms in some specific conditions. Basically, most of the phase-change materials are the ternary GeTe–Sb2Te3 alloys, among which the Ge2Sb2Te5 (GST) alloy is the prototypical PCM material. The GST can be amorphized in hundreds of picoseconds (ps) and recrystallize in tens of nanoseconds (ns). The cycles of reversible transitions can be as high as \(10^{12}\). The temperature of ten-year data retention is about 85 °C.

Since both optical and electrical signals can be controlled by optical or electrical pulses, PCM is suitable for optical or electrical memories. While the rewriteable optical disk-based on PCM (Figure 3e) have been commercialized for a long time (since 1990), the PCM electronic memory are stepping forward. For example, IBM designed a multibit PCM chip in 2016, which can reliably store 3 bits of data per cell after 10⁶ SET–RESET cycles at temperatures as high as 80 °C. Meanwhile, their realization of MLC in PCM is significant to increase the density and reduce the cost. In 2017, Intel and Micron together delivered an Optane product based on 3D X-Point NVM technology which further taps the potentials of PCM (Figure 3f). It was claimed that the 3D X-Point NVM was 1000 times faster than Flash, and 10 times denser than DRAM. Moreover, PCM shows great potentials in the “in-memory computing” and the artificial neuromorphic system (Figure 3g–i). The former one is an apposite technology to break through the “von Neumann bottleneck” and the latter one is the cornerstone of cognitive computing. In fact, many applications of the “in-memory computing,” such as the four basic arithmetic operations (+, −, ×, ÷), the parallel factorization, the fractional division, and the Boolean-logic gates have been realized by PCM. Most recently, Sebastian et al. at IBM demonstrated a high level computational primitive in a PCM chip (Figure 3g). Using one million PCM devices and a machine-learning algorithm, they successfully detected temporal correlations in a data stream. These applications utilize the dynamic and accumulation effects of crystallization or amorphization processes. On the other hand, many artificial neuromorphic systems also have been demonstrated by PCM. Most recently, Tuma et al. at IBM created an artificial neuron using PCM devices (Figure 3h,i). Temporal correlations in data streams were also detected in such a neuron. Interestingly, the randomness of phase-transition dynamics (melt-quench and crystallization processes) mimics the stochastic feature of biological neuron. As such, population-based computing is also...
realized in hundreds of neurons by exploiting their stochastic dynamics. The advantages of PCM for non von Neumann computing originate from its phase-transition dynamics. First, the materials’ property can be changed gradually by electrical/optical pulses and the change is nonvolatile. This enables the computing and storage at the same location, namely, “in-memory computing.” Second, the property variation during phase transitions supplies analog signals, which are similar to the information transmissions in biological brains. Third, the accumulation behavior (the property changes gradually with the number of pulses) and the reversible switch (the property reverses with overmuch pulses) correspond to the integrate-and-fire function. Fourth, the randomness of the phase-transition dynamics corresponds to the stochastic behavior of biological neurons. Since PCM has many other advantages (e.g., high endurance, fast speed, high density, well-studied mechanism, and mature fabrication technology), it is a prospective candidate for the in-memory/neuromorphic computing. In addition, PCM has also been exploited for some other advanced applications, for instance optical components,\textsuperscript{50} thermal emitter,\textsuperscript{51} flexible display,\textsuperscript{52,53} (Figure 3j) and all-optical NVM\textsuperscript{54} (Figure 3k), where the switching of dielectric of phase-change materials is exploited.

However, the main problem for PCM is the relatively high power consumption. Nowadays, this problem is becoming more and more important in presence of big data. In detail, the RESET current of PCM materials (such as GST) is still too large.\textsuperscript{55,56} The root of this problem is the energy loss during the melt-quench amorphization or crystallization process. As shown in Figure 4, thermal diffusion is a primary way of energy loss. In the melting process, the material is heated up to its melting point to break the chemical bonds. This process also costs a lot of energies. When the material is quenched to amorphous states, some energies are stored in the form of internal energies of amorphous states. However, in fact, most of the energies are dissipated via thermal diffusions. Similarly, the recrystallization process requires a longer-time annealing above...
the crystallization temperature. So, thermal diffusion also dissipate energies during the SET operation.

At present, there are mainly two strategies to solve this problem: 1) device-architecture design and 2) phase-change material optimization. The central idea of the device-architecture design is to concentrate energies. This can be achieved by reducing thermal diffusion, managing the energy distribution, or scaling down the area of the programmable bit. For example, using a low thermal-conductivity TaN as thermal barrier close to TiN electrode and GST, the RESET current was reduced by 90%.[57] By locally blocking the interface between PCM material and the electrode using nanostructured SiO$_x$ to reduce the contact area, the switching power was reduced to 1/20. [58] Using ultrathin carbon nanotube as an embedded electrode, the bit dimension was reduced to the size of the carbon-nanotube diameters (1–6 nm). Consequently, the programming current was reduced by two orders of magnitude comparing with popular devices.[59] On the other hand, the development of new phase-change materials is a more attractive way because it can essentially reduce the energy cost for the material itself. If the new materials are combined with the device-architecture design, the power consumption can be further reduced. The development of new materials has two different directions.[60] One direction is to find materials alternative to GST. For example, during the past decades, there emerged several new material candidates for low power consumption including the SiSbTe alloy,[61] the GeCuTe alloy,[62–64] the GeSb alloy,[65] the TiSbTe alloy,[56] the ScSbTe alloy,[66] the C-doped GST alloy,[67,68] the C-doped GeTe alloy,[69,70] the N-doped GST alloy,[71] the N-incorporated GeTe alloy,[72] the O-doped GST alloy,[73] and so on.[66,74] Another direction is the Ge-Sb-Te superlattice/superlattice-like (GST-SL) materials.[75,76] In recent years, GST-SL has drawn the most of attentions in the field because they reduce the energy consumption while they can still maintain other good performances (such as density, speed, endurance, and signal contrast). The structure of superlattice is composed of
and thus has great potentials in high density data storage. However, we note that all the experimental superlattice samples in this article are only expected to be $[(\text{GeTe})_x/(\text{Sb}_2\text{Te}_3)]_n$, according to the designed fabrication process. In fact, the real components of a superlattice sample often deviate its initial design. Figure 5a compares the resistance–current characteristics of GST-SL and GST alloy during switching. While the resistance contrast between SET and RESET states of GST-SL is comparable with that of GST alloy, the SET and RESET currents for GST-SL are about 1/2 of those for GST alloy. Meanwhile, the SET and RESET voltages are also about 1/2 of those for GST alloy. Thus, the power consumption is significantly reduced. The author reported that the power consumptions of SET and RESET operations in the GST-SL device were 11 and 255 pJ, respectively, while those in a GST alloy device were 90 and 375 pJ, respectively. So, the reduction of total energy (including one SET and one RESET process) is about 42%. This performance still remains after $10^6$ cycles which indicates its good stability. In this device, the diameter of the contact area between GST-SL and electrode is about 75 nm while the thickness of GST-SL is about 40 nm. Figure 5b shows the comparison of the size-dependent endurance for GST-alloy (black circles) and GST-SL (red triangles) based devices. While the endurance of GST alloy becomes worse with the decrease of material thickness, the endurance of GST-SL is still as high as $10^9$ when the thickness is scaled down to 15 nm. Similar to the electrical pulse induced transition, the laser pulse induced switching of GST-SL is also faster than that of GST alloy (Figure 5c). Takaara et al. also demonstrated the low power consumption of GST-SL devices.$^{[77,78]}$ For the device reported in ref. [77], the RESET current density of GST-SL (contact area 100 nm) is 3.3 MA cm$^{-2}$, which is 68% lower than that of GST alloy at the same scale. The RESET voltage is also reduced to 1/2. The device is stable after $10^6$ cycles. Moreover, they demonstrated that GST-SL can build MLC. Tai et al. also achieved a PCRAM based on GST-SL, in which the RESET voltage is about 40% lower than that based on GST alloy.$^{[79]}$ Therefore, the advantage of low power consumption of GST-SL is definitely demonstrated.

Comparing with conventional GST alloy, GST-SL also shows distinctive phase-transition behaviors which has been suggested as the origin of the low power consumption. Figure 5d shows the schematic of GST-SL structure. It is composed of GeTe and Sb$_2$Te$_3$ sublattices separated by vdW gaps. Although the impact of these interfaces is not fully understood, it is reasonable to believe that they played critical roles in the devices.$^{[80]}$ Figure 5e–i shows the TEM images and selected area diffraction (SAD) patterns of the RESET state of a GST-SL. The ordered lattices and diffraction patterns can be seen at the adjacent of the electrode after $10^5$ cycles. This characterization suggests that the RESET state of GST-SL is also crystalline which is quite unusual because the RESET states of previously conventional PCM alloys are amorphous. The authors suggested that the switching of GST-SL was an order-to-order transition without amorphization. Therefore, the origin of low energy cost was attributed to the reduced entropic losses during this transition.

2. The Device Characteristics and Switching Behavior of GST-SL

Figure 5a–j shows the performance and structural feature of GST-SL.$^{[76]}$ In this article, we describe the components of a superlattice using the expression of $[(\text{GeTe})_x/(\text{Sb}_2\text{Te}_3)]_n$. Here, the content in “[ ]” is the repeat unit, $n$ is the number of the unit, $x$ and $y$ are the numbers or thicknesses of GeTe and Sb$_2$Te$_3$ layers, respectively.

Figure 4. Schematic for energy loss during the phase transitions of PCM material where a high temperature melt is usually required.
Furthermore, this mechanism does not involve melting, and thus may substantially reduce the energy dissipation as discussed in Figure 4. The authors also suggested that the transition was achieved by Ge atom flipping between covalently and resonantly bonded sites. This proposed process needs less atomic diffusion that may explain the better endurance at nanoscale. In fact, the structural feature and the working mechanism are still under intensive debates which will be discussed in following sections. In short, the performance of low power consumption and the outstanding phase-transition behavior in GST-SL suggest a novel way to control phase transitions in PCM materials and thus provide new opportunities to improve the performance of PCM devices for big data applications.

3. The Atomic Structure of GST-SL

The performance of a PCM device is determined by the property and structure of the phase-change materials. To optimize the devices, determination of the RESET and SET working structures is the first critical step. Despite the very good performance of the GST-SL device, its atomic picture is still under intensive debates. The superlattice is composed of many layer blocks (here, the layer block is defined as the chemically bonded structure between two nearest vdW gaps, such as the Sb2Te3 quintuple layers and GeTe layers which are horizontally stacked with vdW interactions). Their sequence is an important factor to control the properties. Meanwhile, defects usually influence electrical properties significantly in a semiconductor and also play roles in nucleation. As such, in this section we will introduce current popular models for GST-SL and its existing defects including atomic intermixing, stacking fault and vacancy. Their influences will be discussed as well.

3.1. Popular Atomic Models

At present, there are mainly four GST-SL models with different sequences, namely the Kooi, the Petrov, the Inverted Petrov, and the Ferro models. Figure 6a–d shows the atomic pictures of the four models with a popular composition of Ge2Sb2Te5. It should be noted that the pictures of these models in previous reports only showed the bonds marked by the bicolor solid lines in Figure 6. However, according to the interatomic distances, the interactions marked by extra red dashed lines should be also chemical bonds instead of vdW interactions. Petrov model was proposed in 1968 to describe the stable structure of Ge2Sb2Te5. In 2002, Kooi and De...
Hosson proposed a model by considering pure Ge planes located in the middle of the blocks instead of near the vdW gaps (namely, Kooi model).

By alternating the sequence of (GeTe) sublayers, Petrov and Ferro models were also proposed (see Figure 6). In the past few years, these models have been used to describe the structure of the layer blocks in GST-SL.

Density functional theory predicted that the most stable structure at 0 K was the Kooi model (Figure 6a). The GeTe sublayers are incorporated inside the Sb$_2$Te$_3$ sublayers with chemical bonding. The outmost Te atomic layers form vdW gaps. In the Petrov model (Figure 6b), the GeTe layers locate outside Sb$_2$Te$_3$ layers. The Te atomic layers in GeTe also form the vdW gaps. In the Inverted Petrov model (Figure 6c), the sequence of GeTe atomic layers are inversed, compared to the Petrov model, leading to more vdW gaps. The GeTe and Sb$_2$Te$_3$ sublayers in this model are separated. In the Ferro model (Figure 6d), the GeTe bilayer is with a ferroelectric sequence. By ab initio molecular dynamic simulations, Tominaga et al. suggested that the Inverted Petrov model and the Ferro model became more stable when the temperature rose to >500 K. On the other hand, Yu and Robertson suggested that the Ferro model became the most stable one when the temperature rose to >125 K. Their conclusion was draw according to the enthalpy variations against temperature which were calculated from the phonon dispersion spectrum. Perhaps, the conflicting results from different methods are possibly due to the fact that the Ge/Sb intermixing (discussed in the next section) was not perfectly considered in these simulations.

In fact, the layer blocks with different atomic layer numbers (such as 5, 7, 9, 11, 13, 15, and 17) were found in experiments (Figure 7a–e). The layer blocks with 5 atomic layers are the Sb$_2$Te$_3$ quintuple layers. The layer blocks with 7 to 17 layers

**Figure 6.** a–d) Popular atomic models for the [(GeTe)$_2$/Sb$_2$Te$_3$] superlattice. The numbers indicate the interatomic distances. The bicolored solid lines indicate the normal chemical bonds shown in previous reports. The blue dashed lines indicate the vdW interactions. Although the red dashed lines have been missed in the previous models, we suggest to add these lines according to their atomic distances for chemical bonds instead of vdW bonds. (a)–(d) Reproduced with permission. Copyright 2018, IEEE.

**Figure 7.** a) High angle annular dark field-scanning TEM (HAADF-STEM) image of a [(GeTe)$_1$/Sb$_2$Te$_3$)$_{15}$ superlattice. The film was deposited on Sb-passivated Si(111) surfaces, at a substrate temperature of 230 °C, by molecular beam epitaxy (MBE), and then annealed at 400 °C. b–e) HAADF-STEM images of layer blocks with different atomic layers. f) HAADF-STEM intensity profile of a layer block with 11 atomic layers. g) The statistic of the layer blocks with different atomic layers. (a) Reproduced with permission. Copyright 2017, Royal Society of Chemistry. (b)–(f) Reproduced with permission. Copyright 2015, Published by Royal Society of Chemistry under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.
are the GeSbTe blocks with a pseudo-binary composition of \([\text{GeTe}]_2/\text{Sb}_2\text{Te}_3\]. TEM-intensity profile (Figure 7f) reveals that the Sb and Te atoms locate nearby the vdW gaps, which is similar to the case in the Kooi model. Figure 7g shows the change of distribution of these blocks under annealing. The as-grown samples fabricated by molecular beam epitaxy (MBE) at 230 °C mainly contain the 5 quintuple layers. After annealed at high temperatures (300 and 400 °C), the GeSbTe blocks with 7 and 9 atomic layers increased dramatically at the cost of other layer blocks. This result indicates that the individual GeTe and individual Sb$_7$Te$_3$ are incorporated to form GST layer blocks.

Following the design rules of IPCM devices and the one introduced in most models, the GeTe sublattice targets the (GeTe)$_2$, block of 0.7 nm thick. By opposite, the composition and thickness of the Sb-Te sublattice layer can vary depending on the different group and devices. For example, Simpson et al.\cite{76} Taka et al.\cite{77,78} Ohyanagi and Taka\cite{93} and Saito et al.\cite{94} reported a series of GST-SLs with 4 nm Sb$_2$Te$_3$ sublayers, Tominaga et al.\cite{88,95} reported some GST-SLs with 1 nm Sb$_2$Te$_3$ sublayer. Momand et al.\cite{81,92} and Casarin et al.\cite{70} reported some GST-SLs with 3 nm Sb$_2$Te$_3$ sublayers. Wang et al.\cite{96} reported a GST-SL with 6 nm Sb$_2$Te$_3$ sublayers. Moreover, Kalikka et al.\cite{97} and Noë et al.\cite{74} fabricated a series of GST-SLs with thickness-varied Sb$_2$Te$_3$ sublayers (varied from 1 to 8 nm). Also, Zhou et al.\cite{98} reported a series of GST-SLs with thickness-varied Sb$_2$Te$_3$ sublayers (varied from 1 to 4 nm). Kalikka et al. also demonstrated that the laser fluence for switching decreases with the Sb$_2$Te$_3$ thickness. But there still lacks a systematic investigation on the influence of the Sb$_2$Te$_3$ thickness in electrical devices.

In addition, although the GeTe sublayer in GST-SL was often regarded as \([\text{GeTe}]_2\) in previous models (Figure 6), its thickness in samples/devices have been described as 0.7,\cite{81,99} 0.8,\cite{95} 0.9,\cite{88,95} or 1 nm\cite{76–78,93,94} in literatures depending on the publications and groups. Perhaps, this is due to how the GeTe-sublayer thickness was defined or due to difficulty in the thickness control during material growth of GST-SL. In fact, GST-SLs with varied GeTe-sublayer thicknesses have been fabricated and studied.\cite{95,99} Since the GeTe-sublayer thickness can influence the property of GST-SL, it is worth to pay more attentions to this issue.

### 3.2. Atomic Intermixing

The ideal models were considered to be built based on alternatively arranged pure GeTe and Sb$_2$Te$_3$ sublattices. However, high resolution TEM experiments by Momand et al.\cite{81} Casarin et al.\cite{90} and Lotnyk et al.\cite{91} have demonstrated that some Ge and Sb atoms were intermixed in GST-SLs deposited by MBE and pulse laser deposition (PLD) (Figure 8a–h). It should be noted that the IPCM devices were often fabricated by sputtering method.\cite{76–78} Very recently, Kowalczyk et al.\cite{92} argued in favor and definitively stated the intermixing also happened in high quality GST-SLs deposited by sputtering in industrial physical vapor deposition (PVD) equipment. As such, the switching mechanism of IPCM should be revisited by taking into account the Ge/Sb intermixing. Recently, the pioneering authors, including Kolobov et al., also have stated that “STEM studies on epitaxial GST-SL demonstrated that the GeTe blocks were not located between Sb$_2$Te$_3$ quintuple layers, but were incorporated inside the latter, providing a serious challenge to the early explanation.”\cite{100} Obviously, Ge/Sb intermixing should be energetically stable because it can increase the configurational entropy and release the strain by reducing the mismatch between GeTe and Sb$_2$Te$_3$ sublattices. High angle annular dark field-scanning TEM (HAADF-STEM) intensity profiles in Figure 8a–d indicate that the intermixing situation mainly appears in the cation atomic layers adjacent to vdW gaps. This is possibly because vdW gaps can facilitate the atomic migrations. In Figure 8e, the Sb/Ge ratio in the outmost cation layer of GeTe sublattice is 70%:30%, which is in agreement with the ratio of 75%:25% in theoretical calculations.\cite{89} It is worth noting that the superlattice in Figure 8e was prepared by PLD at a relatively low temperature of 140 °C, but the intermixing is still unavoidable.\cite{91} So the authors argued that the intermixing was a chemically driven process rather than purely thermodynamically driven process. A growth model based on the MBE-fabrication process also was proposed to explain the formation of intermixing.\cite{96} There, GeTe was grown on top of Sb$_2$Te$_3$, and then Ge (or Sb) diffused toward the lower (or upper) layers due to existence of concentration gradients. So, the intermixing is not only energetically favorable but also kinetically unavoidable.

In fact, as early as 1998, Karpinsky et al.\cite{101} and Shelimova et al.\cite{103} have proposed that the stable (GeTe)$_2/\text{Sb}_2\text{Te}_3$ compounds exhibit Ge/Sb intermixing in cation planes. Also, several articles confirmed such model of Ge/Sb intermixing in the GST blocks.\cite{103–107} As mentioned above, recent experiments demonstrated again that the intermixing is a general phenomenon in GST-SLs fabricated by MBE, PLD, and sputtering. Therefore, the GST-SL model with Ge/Sb intermixing should be considered and used in simulations or calculations.

The intermixing significantly affects structures of superlattices. It makes the strictly pure GeTe and Sb$_2$Te$_3$ sublattices hard to be fabricated. Although pure GeTe layers have been observed in the middle of a layer block with 13 atomic layers (Figure 8e), they are incorporated inside the layer block with strong bonding rather than vdW interaction. In other words, isolated GeTe layers are still not easily observed at least in the MBE-growth samples. Instead, rhombohedral GeSbTe layer blocks with different atomic layers (such as 5, 7, 9, 11, and 13) are often observed.\cite{81,89–92} The intermixing also affects the property of GST-SL, such as the local distortions,\cite{89} the bond lengths,\cite{90} and the band gaps.\cite{108} Since the intermixing is a general phenomenon in GST-SL, more researches are needed to verify whether or how the intermixing influences the phase transitions of GST-SL. On the other hand, the role of intermixing of GST-SL on the power consumption for PCM is still required to be clarified.

### 3.3. Stacking Faults

In addition to the atomic intermixing, many stacking faults can be also seen in the HAADF-STEM image (Figure 9a). Such stacking faults have been observed in several reports.\cite{78,81,91,92} Figure 9b shows a HAADF-STEM image that contains a stacking fault (in light blue square). These defects always
appear near the vdW gaps and alter the order of stacking sequence. Takaura et al. suggested that the atoms near the vdW gaps were GeTe layers.\(^{[78]}\) They regarded this defect as the sign (indication) of atomic-sequence change between [Ge-Te-Te-Ge] and [Ge-Te-Ge-Te].\(^{[78]}\) By contrast, Momand et al. identified that the atoms near the vdW gaps were Sb and Te atoms (see Figure 9c).\(^{[81,92]}\) Under thermal annealing, the stacking faults in GST-SL can move and thus change the local structures, for instance, the thickness of the layer blocks.\(^{[89,92]}\) It was reported that the stacking faults can also be generated by electron beam irradiations.\(^{[109]}\) Figure 9d,e shows the in situ observation of the generation of a stacking fault in GST-SL. It was found that the stacking fault can either be stable or grow larger or recover to its initial state (i.e., annihilation of stacking fault). Also, the stacking faults can affect the electrical property of GST-SL. For example, Cecchi et al. fabricated [(Sb\(_2\)Te\(_3\))\(_{2}\) nm/(Ge\(_2\)Sb\(_2\)Te\(_5\))\(_{1.7}\) nm]\(_{10}/(Sb\(_2\)Te\(_3\))\(_{2}\) nm superlattices\(^{[110]}\) where the stacking faults may lower the lateral carrier mobility but can increase the carrier concentrations. In addition, similar stacking faults were also observed in hexagonal GST.\(^{[111]}\)

### 3.4. Vacancy

In crystalline GST alloy, vacancy is a common defect that plays important roles in the phase transitions and optoelectronic properties of GST. For example, 1) the vacancies are in favor of atomic motions, thereby can facilitate phase transitions; 2) the vacancies can scatter the carriers and phonons, thus enhance the Joule heating and suppress the thermal conductivity.\(^{[112]}\)
the vacancies annihilate energetically unfavorable antibonding states and thus influence the optical property;[113] 4) the vacancies may supply carriers to enhance the electrical conductivity;[114] 5) the redistribution of vacancies can also change the electrical property. For example, it was found that a vacancy ordering process (accompanied with a structural transition from cubic to trigonal) can lead to an electronic metal-insulator transition (MIT).[115] In the ideal GST-SL models, the vacancies are gathered to form the vdW gaps. However, Momand et al. suggested that randomly distributed vacancies also existed in GST-SL.[92] In their samples, a substantial amount of 9-layers as the Ge2Sb2Te5 layer block were actually identified to be Ge1Sb2Te4 by transmission electron microscopy energy dispersive X-ray spectroscopy (TEM-EDX) using a 50 nm spot. So, they suggested that vacancies also existed inside the layer blocks. Takaura et al. also have fabricated PCRAMs using a \([\text{Ge}_{z}\text{Te}_{1-z}/\text{Sb}_{2}\text{Te}_{3}]_n\) superlattice with \(z < 0.5\), where there existed vacancies in the GeTe layer blocks.[78] As such, the effect of vacancies inside the layer blocks of GST-SL should be taken into account. To definitely determine the role of the vacancies in GST-SL, their distributions and responses to electrical/optical fields are needed.

4. The Switching Mechanism of GST-SL

Understanding the physical origin of the low energy switching is the basis to improve the performance or develop new devices based on superlattices. Although the order-order transition has been widely believed as the possible mechanism, the atomic phase-transition picture is also under intensive debates. This is the very critical studying direction for superlattice-based PCM at the present stage. In this section, we will introduce and discuss some of the most popular mechanisms proposed by the communities.

4.1. Order-to-Order Transition

As mentioned above, one famous proposal for the low power consumption of the superlattice device is the order-to-order transition between the low resistivity SET state and the high resistivity RESET state. During the transition, the entropic losses are expected to be very small due to absence of melting and thus the free energy changes little. This mechanism has been proposed by Simpson et al.[76] as early as when the IPCM was invented (see Figure 5e–i). However, it has been queried by Momand et al.[81] whether the TEM images by Simpson et al. are sufficient to distinct whether the RESET state is partly (down to 2 nm) or entirely crystalline. Many works have been done for trying to clarify the puzzle. At present, there are mainly two proposals for order-to-order transition, i.e., the Petrov-to-Inverted_Petrov transition and the Ferro-to-Inverted_ Petrov transition. Below are the introductions in detail.

Takaura et al. and Ohyanagi et al. have proposed the Petrov-to-Inverted_Petrov transition mechanism.[77,78,82] The \(I-V\) characteristics have been evaluated by nonequilibrium Green functions density functional theory.[116] The current at \(V = 0.5\) volt in the Petrov model is about 10 times larger than that in the Inverted_Petrov model. This result means that
the Inverted_Petrov model is a high resistive state, while the Petrov model is a low resistive state. However, the calculated ON/OFF ratio of the resistances (1/10) is still one order smaller than the value of 1/100 in experimental IPCM (see Figure 5a). In the Petrov model, the atomic layer sequence is -Sb-Te-Ge-Te-vdW-Te-Ge-Te-Sb-Te- (Figure 10a). The Ge-Te layers are in adjacent of the vdW gaps. When the Ge atoms flip into vdW gaps on both sides, the structure turns into an Inverted_Petrov structure with a sequence of -Sb-Te-vdW-Te-Ge-Ge-Te-vdW-Te-Sb-Te- (Figure 10b). Yu and Robertson have proposed a two-step process for this mechanism and calculated the energy barriers by first-principles calculations.\cite{84} The calculated energy barriers for the first step (vertical flipping) and second step (lateral motion) were 2.59 and 0.05 eV, respectively. Such a large energy barrier for the first step may indicate that the vertical flip is still not very easy to happen. Song et al. suggested that the vertical flipping may take place in the sequential form (atoms flip one-by-one) rather than in the concurrent form (atoms flip collectively).\cite{117} Most recently, Nakamura et al. suggested that the transition may also be a one-step process, i.e., the SET state can be the metastable phase after vertical flipping but before lateral motion, because the calculated resistance of this phase is the smallest.\cite{116} Takaura et al. suggested that the driving force for the Ge-atom flipping was electrical field induced charge injection.\cite{77} Nakamura et al. suggested that the driving force may be the current Joule heat rather than the electrical field effect, because the switching in this model is nonpolar.\cite{116}

Tominaga et al. have proposed a Ferro-to-Inverted_Petrov transition (Figure 10c,d).\cite{83} The calculated \(I-V\) characteristic of the Ferro model shows a 10 times larger \(I\) at \(V = 0.2\) volt than that in the Inverted_Petrov model.\cite{116} This result suggests that the Ferro model is a low resistive state and the Inverted_Petrov phase is still a high resistive state. However, the magnitude of signal contrast is still smaller than the experiment.\cite{76} In the Ferro and Inverted_Petrov models, the atomic layer sequences are -Sb-Te-Ge-Te-vdW-Te-Ge-Te-Sb-Te- and -Sb-Te-vdW-Te-Ge-Ge-Te-vdW-Te-Sb-Te-, respectively. So, the transition can be achieved just by one Ge-layer flip. The calculated energy barriers for the vertical flipping (first step) and lateral motion (second step) calculated by Yu and Robertson were 2.56 and 0.39 eV, respectively.\cite{84} Then, the total energy barrier (2.56 + 0.39 = 2.95 eV) for this proposal are also very large. Tominaga et al. suggested that the polarized electrical field should be the driving force for this transition because its switching is in a polar mode.\cite{83,116} Bolotov et al. have studied the voltage-pulse induced switching of electric conductance by scanning probe microscopy and scanning probe lithography.\cite{118} In some areas of their samples, the current responded with a time delay from 0.05 to 10 s after the voltage was applied. They argued that the existence of time delay indicates the switching was achieved by the electrical field effect rather than the current effect.

Kalikka et al. searched the possible structures of GST-SL using a genetic algorithms method.\cite{119} They proposed new structural models that are different from the previous ones (Petrov, Inverted_Petrov, Ferro, and Kooi models). In their models, the GST-SL is composed of stoichiometric Ge\(_3\)Sb\(_2\)Te\(_6\) layer blocks, such as vdW-interaction connected Ge\(_3\)Sb\(_2\)Te\(_6\) and Ge\(_3\)Sb\(_2\)Te\(_4\) layer blocks. They also regard the Ge-atom flipping as their phase-transition mechanism. When one Ge layer in Ge\(_3\)Sb\(_2\)Te\(_6\) flips into the vdW gaps nearby Ge\(_3\)Sb\(_2\)Te\(_4\), the structure turns into two vdW-connected Ge\(_3\)Sb\(_2\)Te\(_5\) layer blocks. In their reports, the calculated energy barriers for this switching are about 1.83–1.99 eV,\cite{119} which may be still relatively high for fast switching.

Obviously, for this order-to-order transition of GST-SL, the Ge-atom flipping is suggested to play the critical role. Hase et al. and Makino et al. studied the transition dynamics in
4.2. Stacking Fault Assisted Metal–Insulator-Transition

In 2017, almost at the same time, Chen et al.\textsuperscript{[125]} (ILU, China) and Kolobov et al. (AIST, Japan)\textsuperscript{[126]} have independently proposed a new mechanism that the metal–insulator switching in GST-SL can be achieved by stacking-fault motions. Generally according to electron counting model\textsuperscript{[127,128]} chemical stoichiometry guarantees the existence of band gaps in various semiconductors. The first-principles calculations of Chen et al. further demonstrated that breaking the local stoichiometry with total composition unchanged can also eliminates the band gap (see Figure 11a–c). Figure 11a shows a model composed of two stoichiometric layer blocks where the upper block is Ge\textsubscript{1}Sb\textsubscript{5}Te\textsubscript{8} and the lower one is Sb\textsubscript{3}Te\textsubscript{4}. Both layer blocks have been observed in experiments. An obvious band gap of this model can be seen in the density of states (DOS) in Figure 11c. Figure 11b shows a model that has the same total composition with the one in Figure 11a (i.e., Ge\textsubscript{1}Sb\textsubscript{5}Te\textsubscript{8}). But it is composed of two nonstoichiometric layer blocks where the upper block is GeSbTe\textsubscript{3} and the lower one is Sb\textsubscript{3}Te\textsubscript{4}. The DOS in Figure 11c demonstrates that the band gap of this model disappears. Interestingly, in experiments it has been observed that stacking faults in GST-SL moved under annealing or electron-beam irradiation\textsuperscript{[91,92,109,125]} Figure 11k,l displays and compares the theoretical model and experimental observation of a stacking fault in GST-SL\textsuperscript{[126]} In fact, the motions of these stacking faults can be realized by the flipping of Sb atomic layers into a neighboring vdw gaps (see another model in Figure 11d–h). If we suppose all the local layer blocks in the as-grown superlattice are stoichiometric, this flipping will change the composition of layer blocks and breaks the local stoichiometry. For example, the local compositions of both the upper and lower layer blocks in Figure 11d are stoichiometric. After finishing the motions of the stacking fault (see Figure 11h), the local compositions of the upper and lower layer blocks become nonstoichiometric Ge\textsubscript{1}Sb\textsubscript{5}Te\textsubscript{8} and Ge\textsubscript{1}Sb\textsubscript{5}Te\textsubscript{6}, respectively. In other words, the local stoichiometry in layer blocks is broken. Indeed, an insulator–metal transition is achieved by such stacking-faults motion (Figure 11i). This transition can lead to prominent change of carrier concentration for information recording. The concentration of thermal excited electrons and holes can be estimated by the integration of the DOS and Fermi–Dirac distribution function

\[
n_n = \frac{1}{V} \int_{E_F}^{\infty} \text{DOS} \cdot f_{FD}(E,T) \, dE
\]

\[
n_p = \frac{1}{V} \int_{-\infty}^{E_F} \text{DOS} \cdot [1 - f_{FD}(E,T)] \, dE
\]

Here, the DOS is the result in Figure 11i, \(V\) is the volume of our model, \(T = 300\) K is the room temperature, and \(f_{FD}(E,T) = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)}\) is the Fermi–Dirac distribution function. The estimated carrier concentration \(n = n_e + n_h\) is plotted in Figure 11j.

Compared to the conventional phase transition between the crystalline and amorphous states, the stacking-fault motions can save the switching energy significantly. In previously proposed Petrov/Ferro–Inverted Petrov transitions, the energy barrier for the Ge-layer flipping is \(>2.56\) eV. However, the energy barrier for the stacking-fault motion is much lower than the Ge-layer flipping mechanisms. The calculated energy barrier for the stacking-fault motion by assistance of vacancy, where the atoms at the boundary of stacking faults move one-by-one, can be as low as 0.5–0.7 eV.\textsuperscript{[89,125]} Since the Sb-atom flipping plays critical roles in the transition, the energy barrier of the flipping can significantly influence the possibility of transition. As has been discussed by Chen et al. in ref. [125], the rate coefficient of flipping can be estimated by the Arrhenius equation

\[
k = A e^{-E_a/k_B T}
\]

For the atom flipping process, \(A\) should be the frequency of the atom vibration along z axis. The frequency of the \(A_{1g}\) phonon in GST-SL is about 3.48 THz.\textsuperscript{[120]} \(K_B\) and \(T\) are the Boltzmann constant and temperature, respectively. \(E_a\) is the activation energy (per atom) of flipping. Suppose that the temperature is close to its melting point (\(\approx 900\) K). The rate coefficient of the flipping with \(E_a = 0.5\) calculated by Equation (3) is about 0.0057 \(\times 10^{12}\) s\(^{-1}\). This result indicates that the atom-flipping time is about 175 \(\times 10^{-12}\) s, which is possible for fast-speed PCM. On the other hand, using the same conditions and an energy barrier of 2.56 eV, the estimated atom-flipping time is about 52 s, which is quite difficult for fast data storage.

Moreover, this mechanism mainly relies on the flipping of Sb atoms instead of Ge atoms. After one Ge layer in a layer block flips to bond with neighboring layer block, the change of the local composition of the layer block is, in fact, GeTe, which does not break the local stoichiometry. By contrast, the flipping of one Sb layer leads to a composition adding or subtracting a SbTe bilayer, which inevitably breaks the local stoichiometry. In other words, the local stoichiometry will still be broken as long as the Sb atoms flip. So, this mechanism can also work in a Ge/Sb intermixed model. However, at the present stage the detailed process of reversibility is still unclear. That will need more investigations in near future.
4.3. Strain Assisted Phase Transition

In 2016, Zhou et al. [98] and Kalikka [97] et al. proposed a strain assisted phase-transition mechanism. The in-plane lattice parameters of Sb$_2$Te$_3$ and GeTe bulks are 4.26 and 4.16 Å, respectively.[97] So, when a GeTe layer is incorporated in a [(GeTe)$_x$(Sb$_2$Te$_3$)$_y$]$_n$ superlattice, it is tensilely strained. The strain will increase with increasing thickness of Sb$_2$Te$_3$ layers. It was reported that the strain can reach a maximum value ($\approx$ 2%) in a superlattice with 4 nm Sb$_2$Te$_3$ layers and 1 nm GeTe layers.[97] With increasing the strain, the switching barrier for the Ge-atom flipping decreases substantially (Figure 12a,b). They proposed that the flipping of Ge atoms to vdW gaps triggered the disordering of GeTe layers. Then the amorphization of GeTe layers can take place with a lower temperature than the melting point and thus reduce the switching energy.

They also performed ab initio molecular dynamics simulation using a Petrov model where the GeTe layers gradually amorphized, but with Sb$_2$Te$_3$ layers unchanged. So, this transition is a partial amorphization which further reduces the energy cost for switching. They tried to support this mechanism by experiments of electrical/laser pulse induced RESET switching.[97,98] For example, in ref. [98] [(Sb$_2$Te$_3$)$_2$ nm/(GeTe)$_1$ nm]$_{13}$ superlattice showed a lower switching voltage and faster switching speed than [(Sb$_2$Te$_3$)$_2$ nm/(GeTe)$_1$ nm]$_{13}$ superlattice. This was explained by the argument that the strain of GeTe in the former

![Diagram](https://example.com/diagram.png)
superlattice is larger than that in the later one. In ref. [97], the laser fluence to trigger of the switch of the GST-SL decreased with the thickness of Sb$_2$Te$_3$ layers which may be related to increasing GeTe strain (see Figure 12c). The threshold fluence of a [(Sb$_2$Te$_3$)$_4$ nm/(GeTe)$_1$ nm]$_8$ superlattice can reduce 23% comparing with that of GST alloy. This mechanism assumes the well separation between GeTe and Sb$_2$Te$_3$ layers in GST-SL. In fact, some experiments[81,90–92] have observed the intermixing between Ge and Sb in GST-SL. Therefore, this mechanism requires to avoid the intermixing as much as possible.

**4.4. Partial Melting or Partial Amorphization**

In conventional PCM devices, the phase transition happens between the crystalline (order) and amorphous (disorder) phases in certain region. So, it is natural to consider whether the order-to-disorder transition in GST-SL can also take place with fewer energies to match the requirement of GST-SL devices. In 2006, Chong et al. achieved lower programing current and faster working speed in PCRAM based on superlattice-like GST.[75] The main reason for the low energy consumption was attributed to the low thermal conductivity. Because a lower thermal conductivity makes it more effectively to utilize the thermal energy, namely the thermal energy is more focused on the small area to melt the materials. For amorphous (RESET) states, the thermal conductivity of superlattice-like GST measured by 3ω method was less than 0.2 W m$^{-1}$ K$^{-1}$,[129] which is comparable with those of as-deposited GST (0.2 W m$^{-1}$ K$^{-1}$),[76,112,129] GeTe (0.23 W m$^{-1}$ K$^{-1}$),[129–131] and Sb$_2$Te$_3$ (0.32 W m$^{-1}$ K$^{-1}$).[129] However, the thermal conductivity of RESET-state GST-SL measured by a pump–probe thermoreflectivity technique was about 0.33 W m$^{-1}$ K$^{-1}$, which is larger than amorphous GST.[76] Therefore, Simpson et al. argued that thermal conductivity cannot explain the low power consumption of SET process for PCM. For SET state, the thermal conductivity of GST-SL measured by coherent phonon spectroscopy was about 1.9 W m$^{-1}$ K$^{-1}$,[132] which is comparable with that of hexagonal GST (1.6 W m$^{-1}$ K$^{-1}$)[113] but is much larger than that of rock-salt GST (0.5 W m$^{-1}$ K$^{-1}$).[112,133] So, the thermal conductivity also cannot explain the low power consumption of RESET process for GST-SL. In addition, the thermal conductivity of SET-state GST-SL is still less than those of crystalline GeTe (3.1 W m$^{-1}$ K$^{-1}$),[110,113] and Sb$_2$Te$_3$ (2.5–4.7 W m$^{-1}$ K$^{-1}$).[114–116] It should be noted that the variation of thermal conductivities of GST may come from different samples and methods. On the other hand, the anisotropy of GST-SL may also have influences on measuring its thermal conductivity.

In addition, the switching current (I) in GST-SL is lower than that in GST alloy, while the resistance (R) of GST-SL is almost the same with that of GST alloy (see Figure 5a). Then the Joule heat (I$^2$R) in GST-SL is also much smaller than that in GST alloy. In summary, the argument of low thermal conductivity may need to be revisited to explain the low power consumption.

In rock-salt GST, the severe phonon scattering caused by the lattice distortions and vacancies is responsible for the extremely low thermal conductivity.[112] Since the disordered vacancies and distortions in GST-SL is less than those in GST alloys, the phonon scattering at interfaces should be responsible for the thermal conductivity of GST-SL. This is supported by the fact that the thermal conductivity of superlattice-like GST generally decreases with the number of interfaces.[129] So, a possible thermal-melting argument is that most of the thermal energies focus on or acts on a limited region, for example, the areas nearby the interfaces. Further explorations on the distribution of Joule heat under an electrical field should be helpful to address this issue.

Generally, there are two types of interfaces in GST-SL. The first type is the vdW gaps which supply spaces for atomic movements and thus possibly serve as a nucleation site for melting. Similar to GST-SL, the recently found vacancy ordered cubic (VOC)-phase GST also has vdW gaps. By ab initio molecular dynamic simulations, Wang et al. have demonstrated that partial melting occurred around vdW interfaces in the VOC.
widely believed that the crystallization in conventional GST alloy is nucleation-dominated, thereby GST in devices often appears as polycrystals. However, the electron diffraction pattern in Figure 5 demonstrates that the GST-SL is a single crystal. As such, the recrystallization in GST-SL may be growth-dominated. That is possibly in favor of partial-melting. Since the partial melting easily occurs at the interfaces, vertical resistance will change a lot but the lateral resistance will change a little. Then, the change of the lateral resistance can serve as a standard to examine this mechanism in near future.

5. Optimization of Superlattice for Advanced Performance

The superlattice is a relatively new material, there are significant opportunities to further optimize its materials and devices, such as improving the crystal quality, adjusting the structures, and searching for better compositions for storage. In this section, we will review the fabrication methods and the influence factors from material structures and compositions on performances of its devices. Possible approaches to optimize this kind of materials will be suggested as well.

5.1. Optimization of Fabrication Methods

The reported superlattice PCRAM devices are mainly fabricated by magnetron sputtering and PVD techniques. For example, Simpson et al., Tominaga et al., and Ohyanagi and Takaura prepared their superlattice materials by helicon-wave RF magnetron sputtering at about 200–250 °C on Si substrates. Takaura et al. fabricated their GST-SL devices using the PVD methods. Kowalczyk et al. also prepared the GST-SLs on Si substrates by sputtering method at 250 °C. Kawamura et al. fabricated the superlattices by the MBE at 230 °C on Si substrates. Since the temperature higher than 220 °C is also favorable for formation of GST alloy (i.e., cancelation of SL) and some defects, Lotnyk et al. prepared the superlattices by the MBE technique with a temperature of 140–145 °C, which is much lower than those with the sputtering and MBE methods. However, the formation of atomic intermixing and stacking faults were still found by the PLD method. As such, the formation of these defects are not only thermally driven but also kinetically or chemically driven.

During the process of crystal growth, the substrate often plays critical roles. For example, the mismatch between crystal and substrate will introduce stress in materials and thus changes the structures and properties. Since bulk Sb$_2$Te$_3$ is a kind of material composed of quintuple layers and vdW gaps, some works investigated the growth of Sb$_2$Te$_3$ on substrate first before to study GST-SL. Kolobov et al. demonstrated that the crystal quality of GeTe films on Si (111) surface (grown by MBE) was better than the films grown on other surfaces. Wang et al. investigated the growth of GeTe thin films on reconstructed Si (111) surfaces. They found that the crystal quality of the films grown on a Sb passivated Si(111)-$(\sqrt{3} \times \sqrt{3})$R30°-Sb surface was improved comparing with those...
grown on a Si(111)-(7 × 7) surface. Similarly, Boschker et al. fabricated Sb2Te3 films on a substrate of Si(111)-(√3 × √3)R30°-Sb (by MBE), where coincidence lattice was formed between Sb2Te3 and Si(111) surface (Figure 14a–c).

In fact, GST-SLs have been grown by MBE using GeTe and Sb2Te3 as starting layers (contacting to a substrate of Si(111)-(√3 × √3)R30°-Sb surface), respectively. The GST-SL starting with Sb2Te3 shows better quality than that starting with GeTe. In addition,

Boschker et al. have also fabricated Sb2Te3 films by MBE on graphene substrates, where coincidence lattices between Sb2Te3 and graphene were formed. Such an attempt promotes more choices of the substrates in GST-SL fabrication. Saito et al. have proposed a self-organized van der Waals epitaxy of Sb2Te3 and GST-SL (Figure 14d–k). During a sputter deposition process, Sb and Te atoms are adsorbed on the substrate.

If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected.

Figure 14. a) Schematic of Sb2Te3 films grown on Si substrate. Angular-resolved reflection high energy electron diffraction intensity for Sb2Te3 on b) Si(111)-(7 × 7) and c) Si(111)-(5√3 × 5√3)R30°-Sb. (a)–(c) Reproduced with permission.[146] Copyright 2014, American Chemical Society. d–k) Schematic of self-organized van der Waals epitaxy model. Reproduced with permission.[145] Copyright 2015, John Wiley and Sons. d) Atomic structure of Sb2Te3 quintuple layers. e) During a sputter deposition process, Sb2Te3 is self-organized on the substrate. f) If the substrate has no selective reactivity with Sb or Te atoms, it is hard to form quintuple layers with large area. g) If a Te layer is selectively deposited on the substrate, a large-area quintuple layer can be expected. h–k) Cross-sectional TEM images of a [(GeTe)1 nm/(Sb2Te3)4 nm]/(Sb2Te3)5 nm superlattice deposited on an amorphous Si (a-Si) substrate. (h–k) is the bright field image, HAADF-STEM mapping, bright field-magnified image, and fast Fourier-transformed (FFT) image at the interface between the film and substrate, respectively. The resistance versus voltage curves of [(GeTe)1 nm/(Sb2Te3)4 nm] superlattice devices during l) RESET and m) SET operations, respectively. (l),(m) Reproduced with permission.[78] Copyright 2014, IEEE. n) Resistance versus current curves during RESET operations in PCM devices based on [(N-doped-GST)x nm/(GST)2 nm]30 nm superlattice, GST alloy, and NGST alloy. o) Endurance of [(N-doped-GST)x2 nm/(GST)2 nm]30 nm superlattice-like and [(GeTe)x/(Sb2Te3)y]50 nm superlattice-like devices. (n),(o) Reproduced with permission.[151] Copyright 2013, AIP Publishing. p) Cross-sectional high resolution TEM of [(Sn10Te90)4 nm/(Sb2Te3)4 nm]9 superlattice film. q) The switching power ratio of a [(Sn10Te90)4 nm/(Sb2Te3)4 nm]9 superlattice device : GST device. (p),(q) Reproduced with permission.[154] Copyright 2013, AIP Publishing.
or Te atoms, it is hard to form quintuple layers with large area (Figure 14f). By contrast, if a Te layer is selectively deposited on the substrate, large-area quintuple layer can be expected (Figure 14g). According to this mechanism, it is the element rather than the orientation of a substrate that plays critical rules. They suggested several promising elements, including Si, P, Ge, Pb, W, and Re, which can selectively react with Sb or Te. This implies the compounds contains these elements may be suitable as substrates. This mechanism is further confirmed by the fact that highly oriented GST-SL can also be prepared on amorphous Si. Figure 14h–k shows the cross-sectional TEM images of a [(GeTe)\textsubscript{1 nm}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{4 nm}]\textsubscript{10}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{5 nm} superlattice deposited on an amorphous Si (a-Si) substrate. The layered structure of GST-SL is clearly observed. In addition, this grown-on-amorphous technique is done by sputtering. That means it is suitable for larger-area and industrial production. Saito et al. further optimized this method by a two-step process to fabricate GST-SL.[148] There, an amorphous Sb\textsubscript{2}Te\textsubscript{3} seed layer was prepared before the deposition of GST-SL. The grain size of the GST-SL by this way can be larger than 200 nm. Recently, Boschker et al. have demonstrated that GST-SL can be also fabricated on amorphous SiO\textsubscript{2} and amorphous carbons by MBE.[149] It is very attractive to prepare GST-SL or other vdW heterostructures on amorphous substrates because it is favorable to integrate the superlattice with traditional CMOS technology.[149] So, more investigations on this issue are worth to be conducted.

5.2. Optimization of Layer Sequence

As discussed above, the GST-SL contains GeTe, Sb\textsubscript{2}Te\textsubscript{3}, Ge-Sb-Te layer blocks, and vdW gaps. As such, the material can have many different sequences because the number of permutation and combination of these components is very large. It is natural to expect that different sequences contribute different properties. For example, the four primary models (Kooi, Petrov, Inverted_Petrov, and Ferro, shown in Figure 6) have different properties in front of the same composition. Zhao et al. found the device with a sandwich sequence [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{4.2 nm}/(GeTe)\textsubscript{3.5 nm}]\textsubscript{10}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{10} has a lower RESET current, a better endurance, and a better thermal stability comparing with the device with a pair sequence [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{4.9 nm}/(GeTe)\textsubscript{3.5 nm}]\textsubscript{10}. Moreover, the thicknesses of these layer blocks and the vdW gaps can also significantly affect the properties. For example, Kalikka et al. reported that switching energy decreased with the increase of Sb\textsubscript{2}Te\textsubscript{3} thickness.[97] Tan et al. reported that the RESET current varied with the number of layers in superlattice-like GST.[151] Yang et al. also reported that the switching temperature varied with the layer thickness.[152] It should be noted that the superlattice-like phase change materials have been demonstrated to possess an improved power consumption as early as 2006.[78] Despite the difference between the superlattice-like GST and GST-SL,[78] both of them possess the advantage of lower-power consumption. As such, the superlattice-like materials should be also paid attentions to. To summarize, the rule to optimize GST-SLs by adjusting their sequences is a promising subject to control and improve superlattice-PCM devices.

5.3. Optimization of Composition

In addition, the composition of the superlattice is readily tunable to further control the properties of the system. At present, the best composition of the superlattice for memory is still unknown.

Many efforts focus on controlling compositions among Ge, Sb, and Te for superlattice. Takaura et al. reported that the [(Ge\textsubscript{3}Te\textsubscript{1−z}Sb\textsubscript{z})\textsubscript{0.5}]/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{0.5} superlattice with z < 0.5 had a much lower switching current than that with z = 0.5 (Figure 141,m) while the superlattice with z > 0.5 did not work.[78] The origin of this phenomenon was suggested to the Ge vacancies which provides spaces for Ge migrations. In our opinion, the existence of the vacancies in GeTe layers could benefit vertical flips of Ge or partial amorphization. Moreover, despite of the non-stoichiometry of this composition, the number of SET and RESET cycles of this device can reach 1 × 10\textsuperscript{4}. To clarify these issues, the arrangement of the vacancies should be identified clearly in future study. Cecchi et al. fabricated a [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{0.5 nm}/(Ge\textsubscript{2}Sb\textsubscript{2}Te\textsubscript{5})\textsubscript{1.7 nm}]\textsubscript{10}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{2 nm} superlattice, which showed a better lateral mobility than the [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{3 nm}/(GeTe)\textsubscript{1 nm}]\textsubscript{15} superlattice.[150] Yang et al. reported a growth-dominant superlattice-like PCM material composed of alternatively GeTe and Sb\textsubscript{2}Te\textsubscript{3}.[152] With a thickness ratio of 1:6.1 for GeTe:Sb\textsubscript{2}Te\textsubscript{3}, the melting point in this material is 353 °C. The SET and RESET current can be as low as 1 and 1.5 mA, respectively. Zhou et al. reported that the [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{2 nm}/(GeTe)\textsubscript{1 nm}]\textsubscript{13} superlattice had a lower SET voltage than the prototypical [(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{2 nm}/(GeTe)\textsubscript{1 nm}]\textsubscript{13} superlattice.[98] Kowalczyk et al. also have demonstrated that the atomic compositions of GST-SL not only affect the crystal quality but also affect the structures (such as the long-range order or orientation) and thus its properties.[99] Moreover, they demonstrated that the compositions and periods of GST-SL can be well controlled by a cosputtering PVD, which is suitable for industrial fabrications.

On the other hand, external elements have also been proposed to dope in GST-SL. Tan et al. reported a superlattice-like PCRAM composed of alternatively arranged nitrogen-doped GST (N-GST) and GST with a single-layer thickness of 4.2 nm.[151] The RESET current of this superlattice is 39% lower than that of bulk GST (Figure 14n). Moreover, the endurance of the N-doped superlattice can reach 10\textsuperscript{5} cycles, which is better than 10\textsuperscript{5} cycles of GST-SL (Figure 14o). Since many elements (such as C, N, O, Si, Ag, W, Bi, Sn) have been demonstrated to improve the performance of conventional bulk GST,[61,67,71,73,74,153–155] similar attempts in GST-SL should be expected.

In addition, other superlattice (instead of GST-SL) have been reported as well. For example, Soeya et al. reported a PCM device based on [(Sn\textsubscript{2}Te\textsubscript{100−x}I\textsubscript{x})\textsubscript{1 nm}/(Sb\textsubscript{2}Te\textsubscript{3})\textsubscript{4 nm}]\textsubscript{9} superlattice (Figure 14p).[156] In this superlattice, the power consumption can be adjusted by changing the ratio of Sn (Figure 14q). Surprisingly, the lowest power consumption of such PCM device can be only 1/2250 of that of bulk GST alloy. Lu et al. demonstrated the multilevel storage for [(GaSb)\textsubscript{4 nm}/(Sb\textsubscript{2}Te\textsubscript{6 nm})\textsubscript{12} superlattice.[157] This material also show fast SET speed (>10 ns) and good endurance (>10\textsuperscript{5} cycles). Zou et al. fabricated a PCM device using [(Ga\textsubscript{5}Sb\textsubscript{9}O\textsubscript{5})\textsubscript{5 nm}/(Sb\textsubscript{4}Te\textsubscript{3})\textsubscript{5}]\textsubscript{15} superlattice.[158] The device shows fast speed (>10 ns) and low RESET power which
is 6.5 times lower than that of bulk GST. Liu et al. fabricated [(SnSb)10 nm/(GeTe)20 nm] and [(SnSb)2 nm/(GeTe)10 nm] superlattices. Their crystallization times both are <5 ns and the ten-year-retention temperatures are 102 and 150 °C, respectively.\[159\] Hu et al. reported a superlattice composed of [(SnSb)3.5 nm/(Ga3Sb)4 nm]5-superlattice device which had a fast speed (9 ns) and low power consumption (0.4 pJ).\[161\] Liu et al. reported a [(GeTe)4 nm/Sb6 nm]superlattice device which had a fast speed (9 ns) and low power consumption (0.4 pJ).\[161\] Liu et al. reported a multilayer [(SnSb)30 nm/(SbSe)20 nm]5 superlattice material in which the fast switching speed (<5 ns) and good thermal stability (ten-year-retention temperature = 122 °C).\[162\] Saito et al. also predicted that the (SiTe)2/(Sb2Te3)4 superlattice may be a possible phase change materials.\[163\] To summarize, the advantages of superlattice or superlattice-like phase change materials do not exclusively rely on the composition of [(GeTe)4/(Sb2Te3)3]x. Searching for new compositions with other element should be a promising way to further optimize the superlattice-based phase change memory.

6. Conclusion, Outlook, and Potential Application

In summary, we systematically review the development and the most recent progresses of the new PCM material candidate, i.e., GST-SL. The significant advantage of GST-SL is the low power consumption which may overcome the current drawback of PCM for application in big data era. Also, GST-SL is a candidate for brain-inspired/neuromorphic computing which offers an opportunity to replace the current von Neumann computing.\[46\] In this review, the properties, structures, switching mechanisms, and possible optimization approaches for PCM superlattice material are discussed and analyzed. The biggest problem in this field should be their unclear structure and switching mechanism. Currently, the proposed switching mechanisms can be divided into two categories. The first one is the traditional (normal) phase transition, i.e., partial melting or strain-assisted transition. This mechanism does not rely on the polarization of applied electric field which is consistent with the most of experiments. However, the reason for the low power consumption is still unclear. The second one is the so-called new (abnormal) mechanism, i.e., order-to-order transition or stacking-fault motion induced MIT. This mechanism is the original proposal in the pioneer work of the IPCM.\[76\] This transition is mainly triggered by Ge/Sb flipping induced by electric field. However, the dependence on the electric-field polarization and how to realize the reversible switching in the second mechanism is still not understood.

Therefore, we suggest that the careful characterization of the atomic structures during switching in device is the key to clarify the puzzle. In other words, the structures of different states (SET/RESET) in real devices should be carefully compared. According to above discussions, the atomic structure of GST-SL is very complicated due to existences of atomic intermixing, stacking faults, and thickness-dependent layer blocks. It is possible that the atomic structures in different devices are quite different. If so, the phase-transition mechanism of GST-SL should be varied. However, the common features of GST-SL, i.e., layered vdW gaps and single-crystal like structure, should be paid attention to.

On the other hand, the optimizations of crystal quality and composition are also important. As for the growth of superlattice, searching for proper substrates is a valuable subject. As for the control of composition, superlattice can extend to other material systems with low power consumption, such as SiSbTe alloy.\[61\] GeCuTe alloy,\[62-64\] GeSb alloy,\[65\] TiSbTe alloy,\[66\] and ScSbTe alloy.\[66\] Their forms in superlattice should have great chances of further improving PCM performance.

PCM superlattices also show potentials in new functions, such as magnetic–electric hybrid storage or topological insulator (TI). Tominaga et al. found a strong response of the magnetic field in a [GeTe]x/(Sb2Te3)36 device.\[88,95,164\] Based on this property, Tominaga and co-workers proposed a hybrid storage device that can be controlled by both electrical and magnetic fields. A magneto-optical Kerr effect in GST-SL has also been reported.\[165-167\] In addition, theoretical calculations by Sa et al. indicate that the GST-SL can be a TI.\[168\] The topological insulating property changes with the thickness of component layer blocks. For example, [(GeTe)4/(Sb2Te3)]x and [(GeTe)3/(Sb2Te3)]x superlattices are not TIs while [(GeTe)2/(Sb2Te3)]x superlattice is a TI.\[83,168,169\] These applications and other explorations in future will make GST-SL has broader prospects in electronic engineering.

Acknowledgements

The work in China was supported by the National Key Research and Development Program of China (Grant No. 2017YFB1104300), National Natural Science Foundation of China (Grant Nos. 11874171, 61590930, 61775077, and 11504368), and 973 Program (Grant No. 2014CB921303). The High-Performance Computing Center (HPCC) at Jilin University for calculation resources is acknowledged.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

big data, GeSbTe superlattice, low power consumption, nonvolatile memory, phase-change memory

Received: May 16, 2018
Revised: July 9, 2018
Published online:
