Andrea Redaelli *Editor*

Phase Change Memory

Device Physics, Reliability and Applications

Phase Change Memory

Andrea Redaelli
Editor

Phase Change Memory

Device Physics, Reliability and Applications

Editor
Andrea Redaelli
Micron Semiconductor Italia S.r.l.
Vimercate, Italy

ISBN 978-3-319-69052-0 ISBN 978-3-319-69053-7 (eBook)
https://doi.org/10.1007/978-3-319-69053-7

Library of Congress Control Number: 2017959302

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my family, my friends, my colleagues,
and to anyone that spends his life looking
for progress of mankind…

"…fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza."
Dante Alighieri

Foreword

Few scientists and engineers would have thought that it would have taken about 50 years to turn the ideas which S.R. Ovshinsky described in 1968 into a viable technology. In his seminal paper "Reversible electrical switching phenomena in disordered structures," some of the foundations of an emerging memory technology were created. Yet, as the present look describes in pleasant depth and detail, it was still a long journey before a mature technology base was developed. The mechanism exploited in phase change memories is the change of electrical resistance upon the transition from the amorphous to the crystalline state. Such a resistivity change upon crystallization is observed in almost all solids. Nevertheless, the requirements the phase change memories have to meet are significantly more demanding. This can be easily seen if just a few of these requirements are listed. For a good memory, we expect the stored information to be secure for long periods of time, i.e., ideally more than 10 years for a nonvolatile memory. This implies that the amorphous state, which is less stable than its crystalline counterpart, is stable at room temperature and slightly above for these 10 years. Yet, we also want to erase amorphous bits, if needed, within a few tens to hundreds of nanoseconds upon applying short heating pulses. Hence, upon heating the amorphous state should crystallize on a nanosecond time scale. Ultimate phase control is hence a prerequisite for a successful memory technology based on phase change materials. However, the requirements for phase change materials go beyond this. A pronounced resistance contrast is needed, too. As discussed in Chap. 2 of the book, this is enabled by the unique transport properties which govern phase change materials. To realize the unusual crystallization kinetics, which were mentioned above, the concept of fragility is introduced and discussed in Chap. 3. Phase change materials are bad glass formers and show an unusual temperature dependence of their viscosity. Fortunately, in recent years the understanding of material properties has advanced considerably, so that even the behavior of memory cells can be simulated self-consistently, as discussed in Chap. 4, enabling a quantitative description of key features. This has also helped studies of memory reliability which focus on data loss by undesired crystallization and ways to avoid it. The two subsequent chapters

describe material characterization and material engineering for device optimization. For a memory technology to be successful, scaling is a major challenge, discussed in detail in Chap. 8. A second major characteristic is power consumption, which is closely related to the device architecture employed. Finally, the architecture of entire PCM arrays is presented and the impact of selectors is discussed. The book closes with discussion of application opportunities and performance improvements enabled by PCMs. Hence, the book is ideally suited for engineers who plan to utilize the potential of PCMs. Yet it is also an interesting testimony to the many challenges that had to be overcome in realizing a memory technology based on a highly unconventional class of solids.

RWTH University, Aachen, Germany Matthias Wuttig

Preface

In the last decades, the use of electronic systems expanded in many areas of the modern society: networking, mobile devices, personal computers, cloud storage, as well as social media are part of everyone's life, becoming strong drivers for the semiconductor market increase. Among the semiconductor components, needed to realize the electronic systems, memories are becoming key elements being today one of the main contributors in defining the overall cost and system performance. In mobile applications, for example, the smartphones and tablets, the memory capability is often used as a key feature to justify the system cost to the costumer. Furthermore, the other key feature that the costumer considers is the duration of the battery that is related to the system power consumption being again strongly impacted by the memory chips. In computer applications, the transition from magnetic hard drive to NAND-based solid-state disk boosted significantly the system performance both in terms of power consumption and speed. For data center server applications (the reader can think, e.g., to search engines and social media servers), the access of large amount of information requires high memory densities operated at very high speed. Today this is done mainly through volatile DRAM chips that need continuous refresh, thus adding very high power consumption on top of the one related to the CPUs.

Up to now, memory demand has been satisfied by the optimization and miniaturization of standard memory technologies, i.e., DRAM and NAND flash. However, with the increasing demand of enhanced cell performance and the reduction of the cell dimension much more difficult than in the past, the usual scaling approach is becoming insufficient. On the one hand, the scaling of DRAM is more and more critical, thus translating in a cost increase; on the other hand, NAND flash technology performance is often poor especially when the application needs a fast communication between the CPU and the memory element. In this framework, new alternative memory technologies have the potential to bring faster data storage and higher read throughput than NAND while still meeting the nonvolatility requirements and providing a significant cost advantage over DRAM.

Furthermore, following the same successful path that we already observed with the introduction of NAND solid-state disks in the computer memory hierarchy, where the nonvolatility and a fast memory access with respect to magnetic disk translated to a significant boosting of system level performances, in the same way this new memory element (often referred in the literature as storage class memory, SCM) can boost the overall system performances. Despite originally considered as a flash memory replacement, today the phase change memory (PCM) has already arose as the ideal candidate to be used as SCM featuring the required intermediate performances, thus being able to fill the gap between NAND and DRAM in the system.

Despite some books already exist that discuss phase change material and their interesting properties, no books have been published focusing on PCM and its device physics. The aim of this book is thus to review and summarize the learning that has been developed on PCM from a device perspective.

In the first chapter, a short introduction of PCM is reported. Starting from a historical review, the manufactured memory chips with different architectures and densities have been reviewed through the years. The basic working principle and the device physics are then discussed pointing out the peculiar nature of this innovative device based on a phase transformation between a crystalline phase and an amorphous one. The transition is thermal in nature and not conventional for the standard operation of microelectronic devices. In the last part of the chapter, the cell performances are compared to the ones of standard memories highlighting advantages and disadvantages.

The second chapter summarizes the fundamental electrical properties of PCM devices in both the amorphous and crystalline states. First, the band structures of crystalline and amorphous phase change materials are studied based on the analysis of thin films of active materials. Then, the device characteristics in a PCM device including conduction and threshold switching phenomena are shown. Finally, the effects of nonuniform resistivity, the transient phenomena after switching, such as recovery and drift, will be illustrated.

The third chapter is devoted in the first part to the insights into the relevant cell thermal parameters. In the second part, the chapter is focused on the discussion of advanced properties involving the phase change materials in terms of fragility behavior and ionic motion effects. The comprehension of all those pieces is of great importance for a successful implementation of PCM through proper engineering of heat generation and thermal resistances as a very important step for power consumption optimization and program latency.

In the development of a new technology, the availability of a tool able to simulate the cell behavior is key for proper device engineering as extensively discussed in Chap. 4. In fact, the main physical ingredients at the basis of PCM operation are uncommon for conventional electronics, and the available commercial tools cannot be thus employed. In this chapter, a numerical model is described that couples the conduction properties of crystalline and amorphous $Ge_2Sb_2Te_5$ with a local nucleation and growth algorithm to account for the phase transition dynamics. The proposed model can simulate three-dimensional PCM devices, and

it is capable of quantitatively reproducing the key features of chalcogenide physics when integrated in an electronic memory cell.

Chapter 5 gives a review of the reliability mechanisms that are key metrics to enable commercialization of PCM. For PCM the primary risk of data loss is undesired crystallization of the amorphous material phase. This process is stochastic in nature and is accelerated by both temperature and electrical bias. Write endurance in PCM is dependent upon both the chalcogenide and the surrounding electrode materials and manufacturing scheme. Due to the potential of the materials to interact at the high programming temperature, these materials must be carefully chosen and integrated to enable a highly reliable cell. With proper device and manufacturing design, the ability to achieve these requirements has been demonstrated in commercial memory products. In addition, as PCM data storage is not based on trapped charge (as in the prevailing floating gate memory devices), it can support applications with radiation-tolerant requirements.

Performances of PCM are intimately linked to the microscopic properties of the phase change materials. The outstanding properties of chalcogenides are reviewed in Chap. 6. The structural and physical properties of the amorphous and crystalline states are illustrated by focusing on two prototypical phase change alloys ($Ge_2Sb_2Te_5$ and GeTe). The origin of the electrical contrast between the amorphous and crystalline states of phase change materials and the nature of the crystallization mechanisms are discussed at microscopic level. The dopant effects on the structure of the amorphous phase and the way they enhance its stability are discussed in detail. Furthermore, the link between the resistance drift of the amorphous state and the structural changes occurring during its ageing is also analyzed.

In Chap. 7, chalcogenide material optimization guidelines for cell performance enhancement are provided, with focus on automotive applications where higher requirement in terms of thermal stability is usually required. Because of some conflicting device performances, for example, the thermal stability of the programmed states and the programming speed, the phase change material needs to be properly engineered. Stoichiometric compounds like $Ge_2Sb_2Te_5$ and GeTe offer short programming times (in the order of a few tens of nanoseconds), but the thermal stability of the programmed states is limited (about 10 years at 100 °C). The addition of light elements into these stoichiometric alloys, for example, nitrogen or carbon, allows decreasing the programming current and extending the temperature range in which the programmed states are stable, thus achieving the specification required for automotive applications (at least 10 years at 150 °C). The use of nonstoichiometric Ge-Sb-Te alloys enables to further extend the performance of the PCM devices. On the one hand, germanium-rich alloys exhibit a high thermal stability of the programmed states. On the other hand, antimony-rich alloys can ensure a very high programming speed suitable for storage class memory applications. Meanwhile, other materials systems, such as Ga-Ge-Sb, have recently demonstrated both a high thermal stability and a high speed, at the expense however of a high programming current.

Chapter 8 will address the basic question of the extent to which phase change memories can be scaled down in size. Fundamental physical limits, along with material properties and device design considerations, all affect the smallest phase change devices that might be achieved and, in this chapter, are discussed in turn. The effects of scaling on key performance parameters are also reported on device switching currents, energies, and speeds. Finally, device performance attributes for production-oriented and research-oriented cells are analyzed and some perspectives for possible future developments provided.

In Chap. 9, the main categories of device architectures that have been studied and realized to exploit the phase change mechanism in electronic devices are discussed. Starting from their first realization at the end of the 1960s, it was apparent that a critical issue for PCM was the programming current (reset current) needed for the phase transition of the chalcogenide material. To minimize the power consumption, several architectures have been proposed, and their characteristics, their advantages, and their weaknesses are here summarized. The element used to transform the current flowing through the device into heat and then temperature (hence the name "heater") has then become a peculiar element of PCM, and the choices for its construction have deep implications not only on the programming current but also on the overall performance of the device. The presence, the position, and the actual geometrical structure of the heater allow the creation of useful classification of PCM architectures that we will exploit in our analysis: self-heating, built-in heater, and remote heater. These categories are then explored highlighting advantages and disadvantages for each of them.

Chapter 10 discusses design architectural solutions and cell management algorithms for a large PCM array. The "large" attribute is fundamental in the architectural definition, as the only cell that is suitable for a high-density array is the one where the selector size matches the size of the memory element. Due to the relatively high PCM cell operating currents, a bipolar junction transistor is the selector of choice for the PCM cell, offering the best current density in a small footprint among other devices. The use of a BJT selector needs a special care compared, for example, to an MOS selector, as the BJT base current is affecting substantially the array biasing. At the highest current density used in program operations, the BJT gain drops and a large portion of the bit line current is transferred to the word line; for this reason, in the chapter the BJT is often referred to as a "diode" even though a fraction of the cell current is still sink by the collector node. On top of the array definition and biasing, the chapter offers an insight on the embedded algorithms used to get the maximum performance from the array and meet the reliability targets for the product. A few notes on embedded ECC and permanent repair storage (fuses) close the subject.

The final chapter (Chap. 11) focuses on how PCM may be used in systems and the potential value it can provide. To accomplish this, first it is necessary to understand the usage of DRAM and NAND in systems today and how any new memory physics (such as PCM) might supplement or replace these existing memories. The value of the new memory is shown as being supplementary to DRAM and NAND, mitigating the deficiencies of these memories to provide both

performance and power improvements to the system. An overview of the features of PCM, both the first generation which went into production and the second generation which did not realize production, is contrasted to the system requirements. The future challenges for PCM and possible future directions are finally reviewed.

Summarizing, this book offers a comprehensive review of all the aspects related to PCM operation starting from the microscopic understanding of chalcogenide material behavior, through the device physics of the basic operation and reliability, up to the cell array management and applications. This book can be considered a good reading for students, professors, and engineers that want to be initiated to this exciting research field, being a condensate of more than 20 years of research done in the best academy and manufacturing companies in the fields.

Vimercate, Italy — Andrea Redaelli

Acknowledgments

I would like to acknowledge all the people that made this project possible. Starting from the very beginning, I would like to thank my Ph.D. thesis tutor, Prof. A. L. Lacaita from Politecnico di Milano, for introducing me to this technology and pushing in me the scientific interest in the physics of those innovative devices. Then, I thank Dot. Roberto Bez who hired me on this program and supported me also when times were tough. I learned from him the right balance between trying to be not only a scientific researcher but also a problem solver with a "make it work" attitude. I thank then Dot. Giorgio Servalli. With him, I had the privilege to see this concept for the first time becoming a real mature product. Finally, I thank Russ Meyer who allowed me to continue studying those devices, supporting the full Italian team and sponsoring this book. Obviously, I thank the high-skilled authors who contributed to this book and all the colleagues who, despite not directly mentioned here, gave a great contribution to the development of this technology.

Contents

1 An Introduction on Phase-Change Memories 1
Agostino Pirovano

2 Electrical Transport in Crystalline and Amorphous Chalcogenide 11
Daniele Ielmini

3 Thermal Model and Remarkable Temperature Effects on the Chalcogenide Alloy 41
Mattia Boniardi

4 Self-Consistent Numerical Model 65
Andrea Redaelli

5 PCM Main Reliability Features 89
Robert Gleixner

6 Structure and Properties of Chalcogenide Materials for PCM 125
Pierre Noé and Françoise Hippert

7 Material Engineering for PCM Device Optimization 181
Véronique Sousa and Gabriele Navarro

8 The Scaling of Phase-Change Memory Materials and Devices 223
Hasan Hayat, Krisztian I. Kohary, and C. David Wright

9 Phase-Change Memory Device Architecture 263
Fabio Pellizzer

10 PCM Array Architecture and Management 285
Corrado Villa

11 PCM Applications and an Outlook to the Future 313
Gregory Atwood

Index 325

Chapter 1
An Introduction on Phase-Change Memories

Agostino Pirovano

1.1 Introduction

The phase-change memory (PCM), also called ovonic unified memory (OUM) or phase-change RAM (PCRAM), is an emerging nonvolatile semiconductor technology based on thermally induced phase transitions of a thin-film chalcogenide material [1]. PCM relies on a resistance change to store data permanently. However, the historical origin of the PCM concept that dates back in the 1970s and the research efforts specifically devoted to this technology by semiconductor industries set them apart from the larger class of emerging memory technologies based on resistance variation and usually referred to as resistive RAM (ReRAM).

The PCM device is essentially a resistor of a thin-film chalcogenide material, usually a Ge–Sb–Te alloy [2] with a low-field resistance that changes by orders of magnitudes, depending on the phase state of the material in the active region (i.e., crystalline or amorphous). In memory operation, cell readout is performed at low bias by sensing the resistance value. Programming requires instead a relatively large current, in order to heat up the chalcogenide material and lead to a thermally induced local phase change. Phase transitions can thus be easily obtained by applying voltage pulses with different amplitudes and with durations in the range of tenths of nanoseconds [3].

Such technology is one of the most promising candidates for next-generation nonvolatile memories (NVM) [4–6] having the potentiality to improve the performance compared to flash – random access time and read throughput (versus NOR flash), write throughput (versus NAND flash), direct write, bit granularity, and endurance – as well as to be scalable beyond flash technology [7].

A. Pirovano (✉)
Micron, Vimercate, Italy
e-mail: apirovan@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_1

1.2 PCM Historical Development

PCM technology relies on the unique electrical properties shown by amorphous chalcogenide alloys. They are compounds with at least one element of the VI group of the chemical table of elements (chalcogens), usually combined with elements of the IV (carbon group elements like silicon and germanium) and V groups (pnictogens). Research on the electronic properties of amorphous chalcogenides dates back to the original work of Ovshinsky in the '50s. From 1958 to 1961, he investigated the chemical and the electrical properties of chalcogenide alloys, demonstrating their electrical behavior as amorphous semiconductors and discovering the existence of reversible switching effects [9].

The unique electrooptical properties of amorphous chalcogenides found a first commercial application in the '70s of the last century for the xerography technology [10, 11], a dry photocopying technique. In this equipment, amorphous selenium was employed for covering the drum used for transferring the ink, exploiting its photoconductive properties that allowed to hold an electrostatic charge in darkness and to conduct it away under light. In those years, the idea of exploiting the huge reversible resistance variation observed in some chalcogenide alloys was proposed for fabricating a solid-state NVM device [10, 12, 13]. It is worth mentioning that the possible exploitation of amorphous chalcogenide for both electrically alterable read-mostly memories and optical storage application based on laser light was largely discussed in [14] by R. G. Neale and J. A. Aseltine. Compared to the first PCM prototype chip described in [12], a second, more efficient, 256-bit array (Fig. 1.1) was proposed where the selecting device, a diode, is placed under the chalcogenide storage element to save additional space and reduce costs. Unfortunately, the chalcogenide alloys used in that first attempt of PCM device was characterized by a slow transition speed and high programming currents, thus making PCM technology not suitable for commercialization.

A significant development in the comprehension of chalcogenide glass physics and for its exploitation in commercial application was achieved starting from the

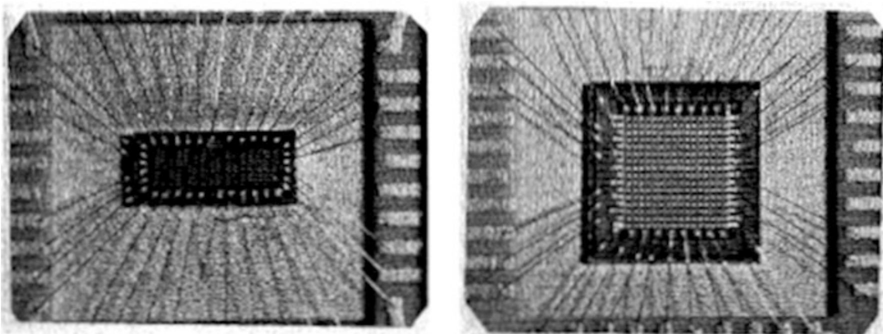


Fig. 1.1 Vertically stacked 256-bit array (left) shown for size comparison with the original 256-bit integrated array (right) [14]

1990s. The Ovshinsky-founded company, Energy Conversion Devices, Inc., developed an amorphous semiconductor alloy made by germanium, antimony, and tellurium, or Ge–Sb–Te (GST), and exploited it in the phase-change technology used in rewritable optical disks [15, 16]. In phase-change optical storage applications, the information storage relies on a laser-induced reversible phase transition [17–19]. A focused laser pulse heats up the GST, which induces the change between the amorphous and the crystalline phases. Since the two phases feature a different reflectance, the optical readout is performed with a low-power probing optical pulse. The GST alloy compositions developed for optical disk at the end of the 1990s are able to switch between the amorphous and the crystalline phase in tens of nanosecond, several orders of magnitude less than the chalcogenide alloys originally employed in the prototype 256-bit PCM device [3, 20, 21]. This achievement renewed the interest of the semiconductor industries on solid-state PCM devices, and several research programs were started around the end of the millennium [2, 3].

In this case, the idea is to induce electrically the phase change in a GST cell and to associate the stored information to the corresponding high and low resistance values. Since both states are stable, no energy is required to keep data stored, resulting in an inherently NVM technology. In order to exploit the PCM concept in solid-state memory devices, the storage element based on the chalcogenide alloy must be coupled with a selecting device that allows to select the individual memory cells for reading and writing. To this purpose, metal–oxide–semiconductor field-effect transistors (MOSFETs) [4], bipolar junction transistors (BJTs) [2], and diodes [22] have been proposed.

The technology development road map of PCM is reported in Fig. 1.2. A test chip fabricated at the 180 nm technology node was used to develop the first demonstrator vehicles and to prove the technology viability [2]. The BJT-selected cell was chosen for the high-performance and high-density application, since the cell size can be $8F^2$ (where F is the minimum cell half-pitch) down to $4F^2$. The MOSFET-selected cell is instead suitable for system-on-chip or embedded application [23], because in spite of the larger cell size (~$20F^2$), the memory integration adds only very few masks to the logic process with a clear cost advantage.

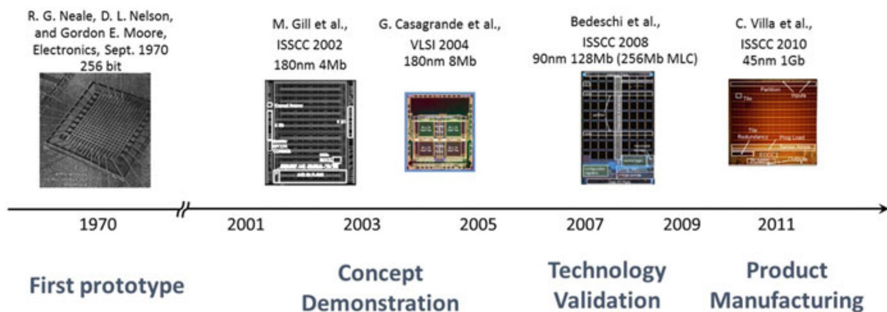


Fig. 1.2 Timeline for the development of PCM technology from the first prototype to the commercial volume production

Successively a 90 nm technology node was developed and a 128 Mb PCM product was commercialized [24]. At this stage of development, PCM was able to deliver legacy performances for NOR replacement as well as new features related to the finer bit granularity and alterability and to the extended endurance capability. Similar preproduction full-spec devices used for the technology development were claimed in [22] with a 512 Mb array at 90 nm technology node. Finally, a 1 Gb PCM product fabricated at the 45 nm technology node with a cell size of $5.5F^2$ was developed, and it has been produced for wireless application [8]. The 45 nm PCM product clearly demonstrates the maturity of the technology [8]. The energy delivered to program a bit is in the order of 10pJ, with a state-of-the-art access time of 85 ns, read throughput of 266 MB/s, and write throughput of 9 MB/s [25]. These peculiar features combined with data retention, single-bit alterability, execution in place, and good cycling performance enable traditional NVM utilizations and novel applications. A similar PCM product has been reported in [26] with a full product based on a 58 nm node technology and with a low-power double-data-rate nonvolatile memory (LPDDR2-NVM) interface. Moreover, PCM is considered one of the best candidates to push to the market the so-called storage-class memory (SCM) [27], a nonvolatile solid-state memory technology that is capable of filling the gap between CPU and disks.

1.3 PCM Working Principle and Device Physics

Some alloys based on the VI group elements (usually referred as chalcogenides) have the interesting characteristic to be stable at room temperature both in the amorphous and in the crystalline phase. In particular, the most promising are the GeSbTe alloys which follow a pseudo-binary composition (between GeTe and Sb_2Te_3), hereafter referred to as GST. The most interesting feature of these alloys is their capability to reversibly switch between a high-resistance amorphous phase and a low-resistance crystalline one in few hundreds of nanoseconds.

The PCM cell is essentially a resistor of a thin-film chalcogenide material with a low-field resistance that changes by orders of magnitudes, depending on the phase state of the GST in the active region. The switch between the two states occurs by means of local temperature increase. Above the critical temperature, the crystal nucleation and growth occur and the material becomes crystalline (set operation). To bring the chalcogenide alloy back to the amorphous state (reset operation), the temperature must be increased above the melting point of hundreds of °C and then very quickly quenched down in order to preserve the disorder and not let the material to crystallize. From an electrical point of view, it is possible to use the Joule effect to reach locally both critical temperatures using the current flow through the material by setting proper voltage pulses (Fig. 1.3).

Depending on the GST phase, the memory device features the electrical characteristics depicted in Fig. 1.3b. The crystalline GST (xGST) behaves as a resistor with a slope mainly determined by the heater resistance itself. The amorphous

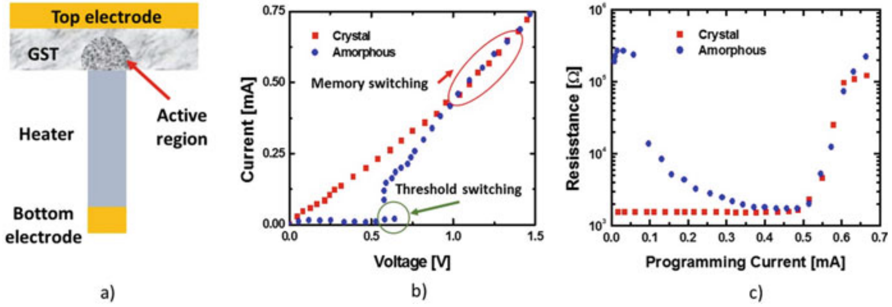


Fig. 1.3 Prototype 180 nm PCM cell (**a**) current voltage characteristics of the cell with the GST in the crystalline and amorphous phase (**b**) and corresponding programming curves (**c**)

phase (aGST) shows an S-shaped negative differential conductance (SNDC), collapsing from a high resistive off-state into a high conductive on-state through the so-called threshold switching effect [28]. Figure 1.3c shows the cell resistance measured after a current pulse of varying amplitude and a duration of 100 ns. Both types of state transitions are shown in the figure: for the reset-to-set transition, the cell was always prepared in the reset state (R $\approx$ 1.5 MΩ) prior to the application of the program pulse. In this case, the programmed resistance steeply decreases for programming currents greater than 100 μA, due to the crystallization of amorphous GST in the active region. The figure also shows the opposite (set-to-reset) transition, where the cell is prepared in the low-resistance state ($R \approx 3\text{k}\Omega$). Here, the cell resistance remains unchanged for moderately small programming currents and then markedly increases for current values greater than 450 μA. This marks the GST melting within the active region. In fact, rapid cooling of the melted chalcogenide along the falling edge of the program pulse results in a final amorphous phase, increasing cell resistance. Note that the falling edge is a critical parameter in the set-to-reset transition: if the falling edge is not fast enough, the melted GST may recrystallize leading to a final state with lower resistance.

From Fig. 1.3b, it is worth noting that there exist two electrically driven switching effects in chalcogenide glasses, as already pointed out in the pioneering work of Ovshinsky [9, 29]. The first one is related to the reversible switching mechanism between an electrical low conductive state and a high conductive one, without any change in the microscopic structure of the material. Since this mechanism requires a threshold voltage and is usually associated with SNDCs, it has been named threshold switching. A second mechanism was also discovered in some chalcogenide compounds, this one related to a thermally activated phase transition of the chalcogenide alloy that can be switched between a high conductive polycrystalline state and a low conductive amorphous one, through proper electrical pulses that cause a Joule heating of the material. Since both the crystalline and the amorphous states are stable, the two phases could be used to store binary information in a nonvolatile memory device, the bit “1” corresponding to the conductive state and the bit “0” to the insulating one. This electrically induced structural

change has been thus named memory switching. It is worth noting that the threshold switching effect is required to get the memory switching mechanism but not vice versa. In fact, an amorphous chalcogenide has to be driven in a highly conductive electrical state through threshold switching to get the transition temperatures without enormous electrical fields [9, 30].

1.4 PCM Device Performance and Applications

PCM represents one of the most promising candidates for next-generation NVM devices [31–33], having the potentiality to improve the performance compared to flash with an exceptional scalability [34]. To exploit the PCM technology for high-performance applications, fast write and read times are mandatory while still preserving good data retention capabilities (Table 1.1).

Compared to other memory technologies, PCM can fill the performance/cost gap between DRAM and NAND, and they can provide a full replacement for NOR flash with a better scalability. In looking at application requirements in comparison with the current memory hierarchy, a very interesting opportunity can be identified for a memory technology with properties in between DRAM and NAND, both in terms of latency and cost. Such a memory system, conventionally defined as storage-class memory (SCM), represents the ideal realm for PCM technologies due to their low-power capabilities. However, stringent requirements in terms of speed and cost must also be met. As reported in Table 1, PCM can deliver much better performance than flash in terms of latency and speed with the attribute of good bit granularity and nonvolatile nature of the storage mechanism.

This perspective PCM technology can be effectively exploited in:

(i) Wireless systems, where LPDDR2-NVM PCM devices have latency in the same order of magnitude as DRAM, could reduce cost and power consumption in the system and can offer reduced application start time and reduced boot time through execute-in-place (XiP) operation.

Table 1.1 Comparison of key attributes among PCM, floating-gate NVM (NOR and NAND flash), and DRAM

Attributes	PCM	NOR	NAND	DRAM
Nonvolatile	Yes	Yes	Yes	No
Clear scaling to	1× nm	45 nm	1× nm	1× nm
Granularity	Small/byte	Large	Large	Small/byte
Erase	No	Yes	Yes	No
Software	Easy	Moderate	Hard	Easy
Power	~Flash	~Flash	~Flash	High
Write bandwidth	1–15+ MB/s	0.5–2 MB/s	10+ MB/s	100+ MB/s
Read latency	50–100 ns	70–100 ns	15−50us	20–80 ns
Endurance	10^6+	10^5	10^4–10^5	Unlimited

(ii) Solid-state storage subsystem to store frequently accessed pages and to store those elements which are more easily managed when manipulated in place. A key advantage is the bit alterable nature of PCM that solves the issue of increased write cycles when the device is full. Higher endurance on PCM addresses the needs of these systems when heavy use is expected.

(iii) PCIe-attached storage arrays, where the fast random read/write operations allow performances much higher than in state-of-the-art flash-based SSD [35].

(iv) Computing platform, exploiting the non-volatility to reduce the power consumption. DRAM, being a volatile memory, consumes watts/GB to simply maintain the contents of the memory. PCM banks can be simply turned off when they are not in use providing reduced power in idle states and decoupling the relationship between density and power consumption.

In order to be able to enter into a well-established memory market, there are key factors that must be fulfilled: (i) match the cost of the existing technology in terms of cell size and process complexity, (ii) find application opportunities optimizing the overall "memory system," and (iii) provide a good perspective in terms of scalability. PCM has been able so far to progress in line with all these requirements.

One of the key challenges that must be faced for industrializing PCM devices is the large programming current needed to melt the cell active region. Such current value may become a limitation in all the applications with a limited budget for power consumption, and it imposes stringent requirements on the current delivered by the memory cell selector integrated in series with the PCM. If the current density delivered by the selecting device is lower than the one required by PCM cell, the selector area must be increased; thus, the size of the cell selecting device becomes the limiting factor for the device density and annihilates the small size advantage of PCM technology. Therefore, reducing the programming current is necessary for achieving both high density and low power consumption of PCM.

Memory cells organized in an array must have a means to select the individual memory cells for reading and writing (RAM operation). The memory cell selection device ensures there is no parasitic current path, the selected cell is writable, and there is adequate read signal-to-noise margin. Ideally, the memory cell selection device has high on-state conductivity and infinite off-state resistance and occupies a small layout area. BJTs and diodes have similar device structures, differing only in the doping of the junctions and the number of contacts [36]. Diodes can have the minimum $4F^2$ layout area. By sharing the base contact with several cells, the footprint of a BJT selector ranges from $8F^2$ [2] to about $5.5F^2$ [8]. The device width of the MOSFET selector is determined by the programming current required. PCM requires a substantial reset programming current that calls for good-quality diode/BJT or wide device width for a MOSFET selector. PN junction diodes with various materials have been studied, including poly-Si, epitaxially grown silicon, doped semiconducting metal oxide, and Ge nanowire: in all these attempts, it has been difficult to find a cell selection diode that simultaneously satisfies the on/off ratio requirement and the on-current required to program the PCM.

1.5 Conclusions

PCM technology represents a serious contender in the emerging NVM arena, being the only technology that has been capable to demonstrate the maturity for large-volume manufacturing at the state-of-the-art technology node. Good electrical performances, superior reliability, and bit granularity can provide a competitive advantage over the mainstream flash technology and make them a good candidate for storage-class memory applications. In order to reduce the cell size and therefore the cost of this technology, two approaches have been deeply investigated in the last decade. First, several alternative cell architectures have been proposed to minimize the programming current. Second, various selecting devices have been experimented to find the optimal solution for different applications and requirements. Finally, MLC and neuromorphic applications represent today the leading edge for the research activities on PCM technology, and they may allow a giant leap for their exploitation in the next years.

With respect to traditional electron device based on the electrical transport properties of silicon, PCM devices exploit several effects that are not usually accounted for in semiconductor devices, at least in solid-state memories. In fact the electrical properties are strongly coupled with the thermal ones, thus requiring to consider the electrothermal physics for all the operating condition of PCM devices. Moreover, for the first time in semiconductor history, the phase-change properties and related physics represent a key topic for device understanding and optimization, thus asking to researchers and engineers operating in this field for new expertise and skills. The aim of this book is to provide a comprehensive treatment of the PCM device physics, including the most recent advances in the modeling of such phenomena and in the comprehension of new chalcogenide material properties, as well as a deep insight in the engineering challenges from the single device to large-sized PCM products.

The electrical properties of amorphous and crystalline chalcogenide alloys have been long debated in the scientific community due to the peculiar features that are unique for these alloys. Chapter 2 reports a complete review of the conduction mechanisms responsible for the subthreshold behavior in chalcogenide alloy as well as the key ingredients to explain the threshold switching phenomenon. Thermal properties and phase-change dynamics are the topic for Chap. 3, with a focus on the phenomena involved in the transition from the amorphous to the crystalline phase. Crystallization dynamics represents the slowest mechanism in PCM operation, and a deep analysis of the crystal nucleation and growth mechanisms is mandatory for the optimization of the device performance. Electrical, thermal, and phase-change physics are then combined in a self-consistent environment described in Chap. 4 and that allow the reader to understand the cross-links between the various mechanisms involved in PCM operations. Apart from the exceptional performance in terms of read and write speed, a key attribute for nonvolatile memories is their reliability to keep the data stored unaltered for a very long time. The physical mechanisms that are responsible for a premature data loss and for a poor device

functionality are deeply discussed in Chap. 5. As mentioned in the historical overview, the optimization of the chalcogenide alloy composition has been the key ingredient for the commercial success of rewritable optical storage and for boosting the renewed interest in solid-state device applications. This has been the first step for a continuous improvement of the device performance based on the development of better and better chalcogenide alloys that are suited for different applications. The scope of Chaps. 6 and 7 is to provide a deep insight in material engineering for optimizing the PCM device performance. Chapter 8 is dedicated to the PCM scaling capabilities. Starting from the fundamental physical limitations for the phase-change mechanism, the ultimate limits for the PCM cell functionality are explored and discussed. To complete our journey in PCM technology, the last three chapters of the book deliver to the reader an unprecedented overview of the PCM cell design strategies, the array architectures, and the specific management techniques adopted for mitigating issues and for exploiting the unique opportunities offered by PCM, with a final outlook in the future of phase-change memory technology and their applications.

References

1. A. Pirovano, A.L. Lacaita, A. Benvenuti, F. Pellizzer, R. Bez, IEEE Trans. Electron Devices **51**, 452–459 (2004)
2. F. Pellizzer, A. Pirovano, F. Ottogalli, M. Magistretti, M. Scaravaggi, P. Zuliani, M. Tosi, A. Benvenuti, P. Besana, S. Cadeo, T. Marangon, R. Morandi, R. Piva, A. Spandre, R. Zonca, A. Modelli, E. Varesi, T. Lowrey, A. Lacaita, G. Casagrande, P. Cappelletti, R. Bez, in *Proc. Symp. VLSI Technology,* 2004
3. S. Lai, T. Lowrey, in *Proc. IEEE Int. Electron Devices Meeting*, pp. 803–807, 2001
4. S. Lai, in *Proc. IEEE Int. Electron Devices Meeting*, p. 255–259, 2003
5. Y. N. Hwang, S. H. Lee, S. J. Ahn, S. Y. Lee, K. C. Ryoo, H. S. Hong, H. C. Koo, F. Yeung, J. H. Oh, H. J. Kim, W. C. Jeong, J. H. Park, H. Horii, Y. H. Ha, J. H. Yi, G. H. Koh, G. T. Jeong, H. S. Jeong, K. Kim, in *Proc. IEEE Int. Electron Devices Meeting*, pp. 37.1.1–37.1.4, 2003
6. Y. N. Hwang, J. S. Hong, S. H. Lee, S. J. Ahn, G. T. Jeong, G. H. Koh, J. H. Oh, H. J. Kim, W. C. Jeong, S. Y. Lee, J. H. Park, K. C. Ryoo, H. Horii, Y. H. Ha, J. H. Yi, W. Y. Cho, Y. T. Kim, K. H. Lee, S. H. Joo, S. O. Park, U. I. Chung, H. S. Jeong, Kinam Kim, in *Proc. Symp. VLSI Technology*, pp. 699–702, 2003
7. A. Pirovano, A. L. Lacaita, A. Benvenuti, F. Pellizzer, S. Hudgens, R. Bez, in *Proc. IEEE Int. Electron Devices Meeting*, pp. 29.6.1–29.6.4, 2003
8. G. Servalli, in *Proc. IEEE Int. Electron Devices Meeting*, pp. 1–4, 2009
9. S.R. Ovshinsky, Phys. Rev. Lett. **21**(20), 1450–1453 (1968)
10. Fundamentals of Amorphous Semiconductors – National Academy of Sciences, Washington DC, 1972
11. M.D. Tabak, S.W. Ing, M.E. Scharfe, IEEE Trans. Electron Devices **20**, 132 (1973)
12. R.G. Neale, D.L. Nelson, G.E. Moore, Electronics **43**(20), 56 (1970)
13. G.V. Bunton, R.M. Quilliam, IEEE Trans. Electron Devices **20**, 140 (1973)
14. R.G. Neale, J.A. Aseltine, IEEE Trans. Electron Devices **20**, 195 (1973)
15. M. Chen, K. Rubin, R. Barton, Appl. Phys. Lett. **49**, 502 (1986)

16. S. Tyson, G. Wicker, T. Lowrey, S. Hudgens, K. Hunt, in *Aerospace Conference Proceedings*, (unpublished), Vol. 5, p. 385, 2000
17. N. Yamada, E. Ohno, K. Nishiuchi, N. Akahira, M. Takao, J. Appl. Phys. **69**, 2849 (1991)
18. J. Coombs, A. Jongenelis, W. van Es-Spiekman, B. Jacobs, J. Appl. Phys. **78**, 4906 (1995)
19. J. Coombs, A. Jongenelis, W. van Es-Spiekman, B. Jacobs, J. Appl. Phys. **78**, 4918 (1995)
20. S. Ovshinsky, Proc. IEEE CAS **1**, 33 (1998)
21. M. Gill, T. Lowrey, J. Park, in *Proc. of ISSCC Tech. Dig.*, pp. 202–459, 2002
22. J. H. Oh, J. H. Park, Y. S. Lim, H. S. Lim, Y. T. Oh, J. S. Kim, J. M. Shin, J. H. Park, Y. J. Song, K. C. Ryoo, D. W. Lim, S. S. Park, J. I. Kim, J. H. Kim, J. Yu, F. Yeung, C. W. Jeong, J. H. Kong, D. H. Kang, G. H. Koh, G. T. Jeong, H. S. Jeong, K. Kim, in *Proc. IEEE Int. Electron Devices Meeting*, 2006
23. F. Ottogalli, A. Pirovano, F. Pellizzer, M. Tosi, P. Zuliani, P. Bonetalli, R. Bez, in *Proc. ESSDERC 04*, pp. 293–296, 2004
24. F. Bedeschi, R. Fackenthal, C. Resta, E. M. Donze, M. Jagasivamani, E. Buda, F. Pellizzer, D. Chow, A. Cabrini, G. M. A. Calvi, R. Faravelli, A. Fantini, G. Torelli, D. Mills, R. Gastaldi, G. Casagrande, in *Proc. of ISSCC Tech. Dig.*, p. 428, 2008
25. Micron Press Release, Micron Announces Availability of Phase Change Memory for Mobile Devices, July 18, 2012
26. C. Villa, D. Mills, G. Barkley, H. Giduturi, S. Schippers, D. Vimercati, in *Proc. of ISSCC Tech. Dig.*, pp. 270–271, 2010
27. H. Chung et al., in *Proc. of ISSCC Tech. Dig.*, p. 500, 2011
28. R.F. Freitas, W.W. Wilcke, IBM J. Res. Dev. **52**(4/5), 439–448 (2008)
29. D. Adler, H.K. Henisch, S.D. Mott, Rev. Mod. Phys. **50**, 209 (1978)
30. S.R. Ovshinsky, I.M. Ovshinsky, Mater. Res. Bull. **5**, 681 (1970)
31. K.E. Van Landingham, IEEE Trans. Electron Devices **20**, 178 (1973)
32. A. Fazio, IEEE IEDM Tech. Dig. pp. 267–270, 1999
33. S. Keeney, IEEE IEDM Tech. Dig. pp. 2.5.1–2.5.4, 2001
34. G. Servalli et al., IEEE IEDM Tech. Dig., pp. 2.5.1–2.5.4, 2005
35. H. Hu et al., Semiconductor Technology International Conference (CSTIC), China 2015
36. Moneta and Onyx: Very Fast SSDs – http://nvsl.ucsd.edu/moneta/

Chapter 2
Electrical Transport in Crystalline and Amorphous Chalcogenide

Daniele Ielmini

2.1 Phase-Dependent Band Structure

Thin films represent a valuable vehicle for characterizing the electrical properties of materials, e.g., the electrical resistivity in different phases and temperature, as well as for carrying out optical investigations (e.g., by photon absorption or reflection) of the band structure for various structural phases. Some properties of the band structure, e.g., the energy gap, its direct/indirect character, the presence of tails of localized states, etc., can in fact be characterized almost directly from optical experiments, thus providing a key input into the physical mechanisms for conduction that are responsible for the electrical properties of the materials.

Most of the results shown in the following refer to $Ge_2Sb_2Te_5$ (GST), which is by far the most widely studied phase change material for optical and electrical storage since the seminal work of Yamada et al. [6]. In fact, GST features a unique combination of excellent properties, such as thermodynamic stability, within the Ge-Sb-Te ternary phase diagram, relatively low melting temperature ($T_m = 620\,°C$), fast crystallization (in the range of 100 ns for sufficiently high temperatures), and good glass-forming ability, in that the maximum quenching time to achieve an amorphous phase by cooling a liquid phase is in the range of about 50 ns [7], thus easily achievable within a nanoscale PCM device connected to a memory circuit. For these reasons, GST is the most popular material used in optical data storage, such as rewritable CD and DVD, and in PCM devices [8–10].

D. Ielmini (✉)
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy
e-mail: daniele.ielmini@polimi.it

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_2

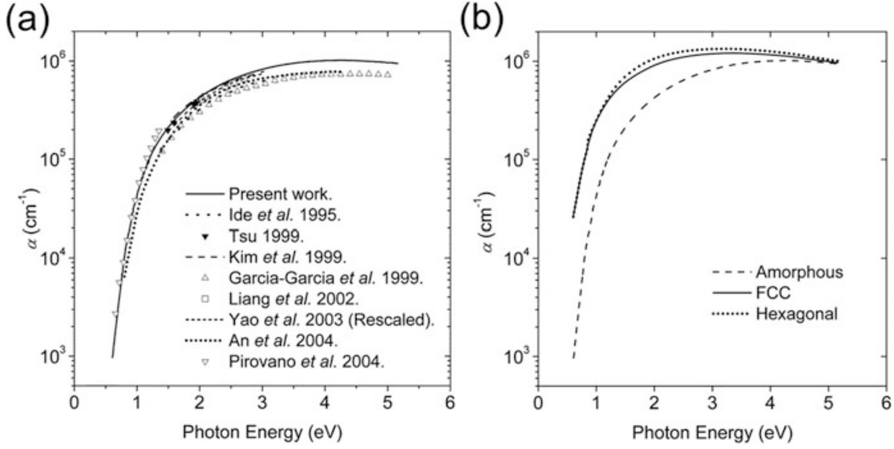


Fig. 2.1 Measured absorption coefficient α as a function of photon energy for amorphous GST (**a**) and comparison between crystalline and amorphous GST (**b**) Reprinted with permission from [11] Copyright (2005) AIP Publishing LLC

2.1.1 Optical Characterization Studies

Figure 2.1a shows the measured absorption coefficient α of GST in the amorphous phase as a function of photon energy $h\nu$ [11]. Data in the figure were obtained from ellipsometry analysis or optical transmission/reflection studies and comparative results obtained from various authors in the literature [11–19], which all show a substantial agreement. Small differences might be attributed to a slightly different stoichiometric composition and/or degree of film oxidation, resulting from the different techniques of deposition. The steep drop of α for decreasing optical energies indicates the band edge. To carefully determine the optical band gap $E_{g,opt}$, data were fitted to the Tauc equation [20, 21]:

$$\alpha h\nu = C\left(h\nu - E_{g,\text{opt}}\right)^r, \tag{2.1}$$

where C and r are constants. Using an exponent $r = 2$, which adequately describes optical electron-hole generation in an amorphous semiconductor where the momentum k is not conserved in the transition [21], and extrapolating to the abscissa, a band gap $E_{g,opt} = 0.7$ eV could be evaluated [11, 19]. More careful analysis indicates a marked deviation from the slope $r = 2$ at low energies below about 0.9 eV, which reveals the presence of exponentially decaying Urbach tails due to localized states. These states are generally attributed to disorder in the amorphous state, according to the Anderson theory of disorder-induced localization of states in the band gap [20]. Such localized states, which are found to belong to a Urbach tail of about 80 meV, have a key role in conduction and switching in the amorphous phase. Similar results for the absorption spectrum and band structure were obtained by subsequent studies [22]. Absorption data were also compared with spectral photoconductivity, measuring the electrical current under illumination of monochromatic

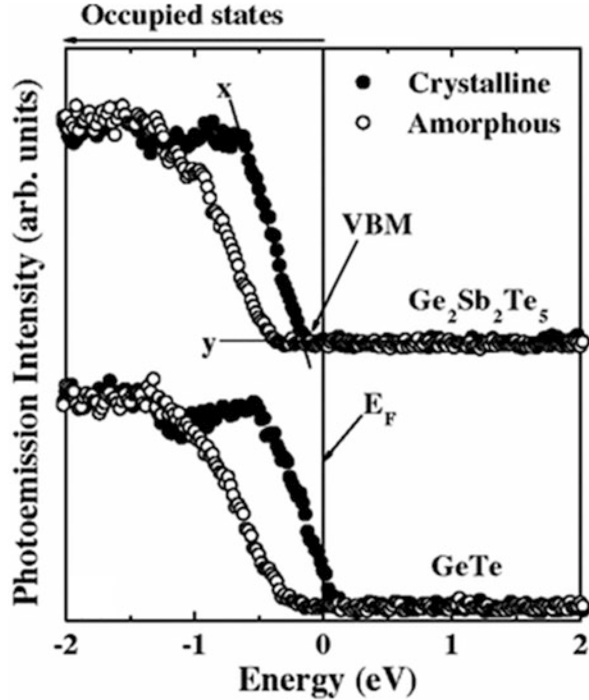


Fig. 2.2 HXPES spectra for GST (top) and GeTe (bottom) in the amorphous and crystalline phases. Reprinted with permission from [23] Copyright (2007) American Physical Society

light at a given photon energy $h\nu$, indicating a substantial agreement and confirming the band structure picture for amorphous GST.

Figure 2.1b shows the comparison of measured α for GST in the crystalline and amorphous phases. Two crystalline phases are considered for GST, namely, the metastable face-centered cubic (fcc) structure and the stable hexagonal structure [6]. The measured α shows a negligible difference between the two phases and a substantially larger value for crystalline phases compared to the amorphous state. This can be understood by a shift of the absorption spectrum to lower photon energies due to a smaller energy gap. Linear fitting in the Tauc plot with $r = 2$ and extrapolation to zero absorption results in an estimated optical energy gap of 0.5 eV [11], in substantial agreement with other reports [14, 15, 17, 19].

Understanding the electrical properties of GST also requires the assessment of the position of the Fermi level, which determines the population of the density of states in the conduction and valence band. To this purpose, Fig. 2.2 shows the experimental results for the hard x-ray photoemission spectroscopy (HXPES) of GST and GeTe [23]. HXPES is a valuable technique to investigate the valence band structure of electronic materials and allows to extract the density of states in the valence band with respect to the Fermi level. The figure shows the HXPES spectra for the amorphous and crystalline phases, as a function of energy, where $E = 0$, represents the equilibrium Fermi energy. The valence band edge is generally closer to the Fermi level for the crystalline phase compared to the amorphous phase,

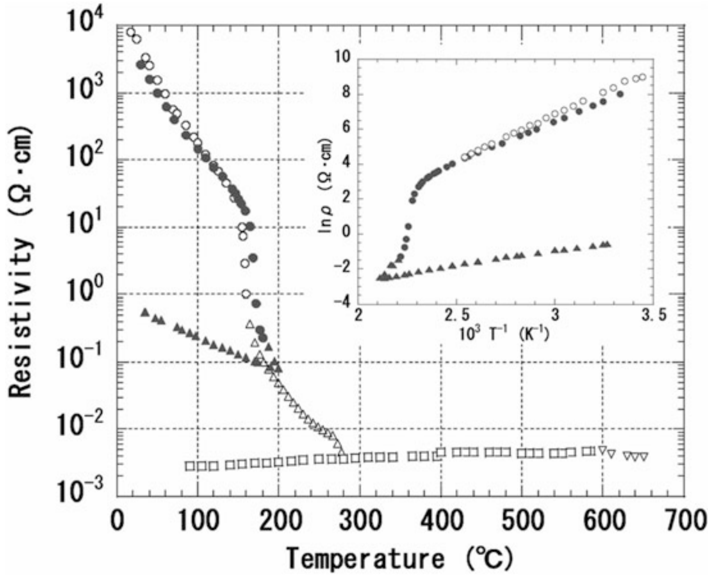


Fig. 2.3 Measured resistivity as a function of temperature for GST in the amorphous and crystalline phases. The film thickness was 110 nm (open symbol) and 510 nm (solid symbol). The inset shows the same data reported in the Arrhenius plot. Reprinted with permission from [22] Copyright (2005) The Japan Society of Applied Physics

supporting the larger conductivity in the crystalline phase which is the basis of the storage concept for PCM devices. For GST, the difference between E_F and the valence band edge in the amorphous phase is about 0.4 eV, i.e., about half of the band gap, indicating that amorphous GST behaves as an intrinsic semiconductor where the Fermi level is pinned at midgap. In the crystalline (fcc) phase, GST shows a deeper valence band compared to GeTe, where E_F lies almost in correspondence of the valence band edge, thus suggesting a degenerate doping concentration and almost metallic conduction. Doping of p-type character is generally attributed to vacancies, which can reach an atomic concentration of about 20% in GST due to vacancies of the Ge/Sb sub-lattice [19, 24], mostly surrounded by Te neighbors [25].

2.1.2 Electrical Characterization Studies

Electrical measurements allow a more direct investigation of carrier transport properties, such as the type of majority carriers and their concentration and mobility. Figure 2.3 shows the measured resistivity as a function of temperature for GST in the amorphous and crystalline phase. In these measurements, the as-grown amorphous GST film was characterized at increasing temperature from 20 to 650 °C, thus above the melting temperature. At relatively low temperature below

150 °C, the amorphous resistivity shows a regular decrease, which can be attributed to the temperature activation of transport in the intrinsic amorphous semiconductor, most probably arising from thermal generation of carriers from the valence to the conduction band with activation energy equal to half of the band gap. The activation energy can be estimated from the slope of resistivity in the Arrhenius plot shown in the inset. Inspection of the Arrhenius plot indicates an activation energy of about 0.45 eV in close agreement with the difference between E_F and the valence band edge in Fig. 2.2. As the temperature reaches about 170 °C, the resistivity shows a marked drop, thus revealing thermally induced crystallization to the fcc phase. If the temperature is decreased soon after crystallization, resistivity increases for decreasing temperature but with a smaller slope compared to the amorphous phase. The Arrhenius plot shows an activation energy of about 0.14 eV, which may be attributed to the energy difference between E_F and the valence band edge for fcc crystalline GST in Fig. 2.2. After the steep region of crystallization, resistivity shows a further decrease with a shallower slope, probably associated with the nucleation and growth of hexagonal grains within the fcc metastable phase. Exceeding the melting temperature does not result in major changes of resistivity, suggesting that the conduction properties of liquid GST are similar to those of the hexagonal phase, which behaves as a degenerate semiconductor, or a metal, given the increase of resistivity with temperature.

Additional insight into the transport properties of amorphous and crystalline GST can be gained by more sophisticated experiments, such as Hall conductivity measurements. Figure 2.4 shows the measured $1/(\sigma R_H)$, namely, the inverse product of electrical conductivity s and Hall coefficient R_H, for the hexagonal and the fcc phases of GST [11, 26]. Data indicate a p-type doping, consistent with the large concentration of Ge and Sb vacancies. An estimate of hole concentration can be derived from the inverse product $1/(\sigma R_H)$, which is in the range of few 10^{20} cm^{-3} at room temperature. A larger hole concentration is found for the hexagonal phase compared to the fcc phase, in agreement with the lower resistivity in Fig. 2.3. Hall measurements are generally extremely difficult for the amorphous phase, due to the extremely low current. However, a p-type character of conduction was assessed for the amorphous phase via Seebeck analysis [22]. This supports the picture of slightly p-type amorphous GST, although its main character remains that of an intrinsic material with very poor concentration of carriers in the extended states. This remains the main origin of the resistivity ratio and optical contrast at the basis of electrical and optical phase change memories, respectively.

Figure 2.5 shows the schematic band structure for the crystalline (top) and the amorphous phases (bottom) of GST [27]. The crystalline phase is a degenerate semiconductor with relatively narrow band gap, where the Fermi level lies close to the valence band maximum due to vacancy-induced self-doping. The amorphous phase appears as an intrinsic semiconductor with the Fermi level close to midgap, as a result of defect-induced pinning. Defects may consist of distorted bonds, wrong bonds (e.g., Ge-Ge bonds in GST), and dangling bonds, which arise due to the disordered structure. Defects also provide localized states in the forbidden gap for capture and emission of carries to/from the band of extended states [28–30]. Although

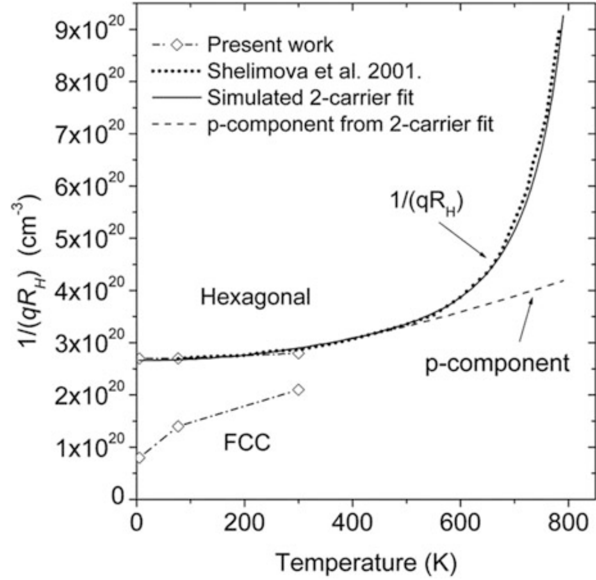


Fig. 2.4 Measured $1/(\sigma R_{\mathrm{H}})$, namely, the inverse product of electrical conductivity σ and Hall coefficient R_{H}, for the hexagonal and the fcc phases of GST [11, 26]. Data indicate a p-type doping, consistently with the large concentration of Ge and Sb vacancies. Reprinted with permission from [22] Copyright (2005) AIP Publishing LLC

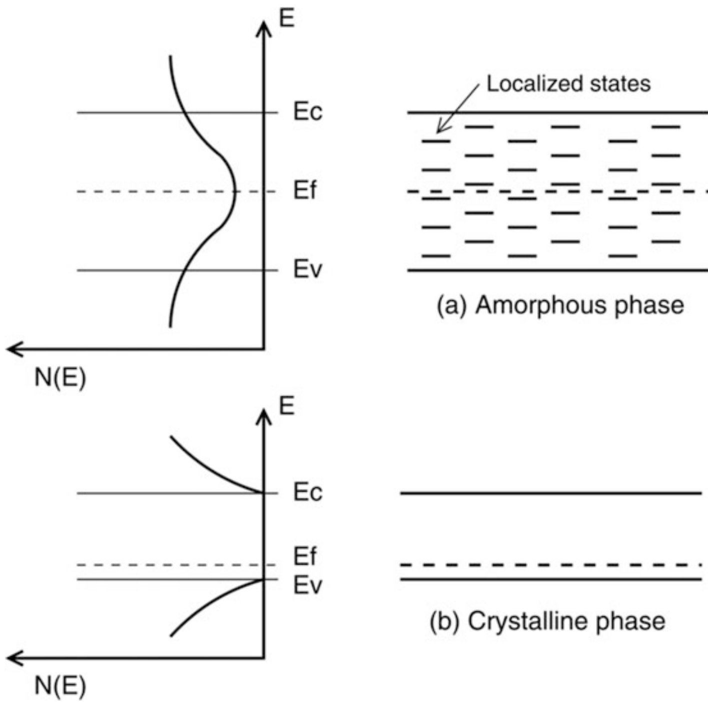


Fig. 2.5 Band structure of the crystalline (**a, top**) and amorphous phases (**b, bottom**) of GST. Reprinted with permission from [27]. Copyright (2008) IEEE

these localized states are preferentially located close to the band edges in the form of Urbach tails, a finite amount of localized state density might take place at the Fermi level deep in the band gap. The different position of the Fermi level eventually accounts for the different electrical resistivity in the two phases at the origin of the PCM storage concept.

2.2 Conduction Characteristics of PCM Devices

While thin films remain the preferred vehicle for basic understanding and exploration of novel materials, the PCM device characteristics are the source of the main information regarding the device and its electrical parameters. For instance, important device properties such as the resistivity of the amorphous phase obtained by quenching from the melt and the threshold switching voltage cannot be obtained from a planar thin film analysis due to the limited electric field and temperature in these experiments.

A PCM device generally consists of a layer of phase change material, such as GST, between two electrical contacts. Figure 2.6 shows the TEM (transmission electron microscope) picture of a PCM device showing a GST film connected to a narrow bottom contact, called heater, and a top thin-film contact [31]. The heater is contacted to one or more select transistors by a metal plug, usually made of tungsten. The select transistor can be a field-effect transistor (FET) as in Fig. 2.6 [31], a bipolar junction transistor (BJT) [10], or a diode-like two-terminal select device [32–36], depending on the required device density in the memory array. In fact, FETs tend to have a larger area due to their planar geometry, thus allowing a relatively small device density. On the other hand, two-terminal selectors, such as polycrystalline silicon p-n junctions [32], ovonic threshold switch (OTS) elements [33], or mixed ionic-electronic conduction (MIEC) devices [34], provide the smallest device area and allow for achieving cross-point arrays with maximum memory density. A detailed review of PCM cells reported so far in the literature is given in Chap. 9.

During the application of a programming pulse, the current flowing in the confined bottom electrode causes Joule heating within the GST volume adjacent to the bottom electrode. For sufficiently intense Joule heating, the GST temperature can exceed the melting point, thus causing local transition to the liquid phase. If the decay time of the applied pulse is short enough, typically below 30 ns [7], the liquid phase is quenched into an amorphous phase, thus inducing reset transition of the PCM. Note that the amorphous volume generally reflects the liquid volume during the reset pulse and thus increases with the amplitude of the voltage pulse [35–38]. This enables multilevel programming of the PCM, where different resistance levels correspond to different thickness of the amorphous volume in the GST layer. On the other hand, the set state corresponds to a wholly crystalline phase, typically consisting of polycrystalline fcc structure.

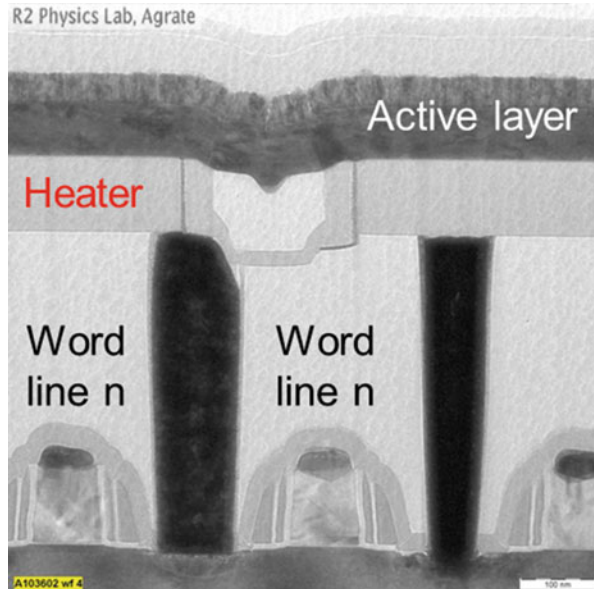


Fig. 2.6 TEM picture of a PCM device with confined bottom electrode, or heater, and non-confined thin-film GST and top electrode. Each PCM device is connected to two transistors acting as select elements for electrical access within the memory array. Reprinted with permission from [31]. Copyright (2013) IEEE

2.2.1 I-V Curves for Set and Reset States

Figure 2.7 shows the measured current-voltage (I-V) curves for a PCM device with confined bottom electrode [10, 39]. In the measured device, the GST film was deposited by RF sputtering [40]. Top and bottom contacts were consisting of sputtered TiN, where a Ti adhesion layer was inserted between the GST film and the top electrode to minimize the contact interface resistance. The width of the BE was confined to 50 nm via conventional lithography, whereas the thickness was 5 nm in the orthogonal direction, thus resulting in a bottom contact area of the PCM device of about 250 nm^2. The GST layer had a thickness of 35 nm. Set and reset states in Fig. 2.7 indicate strongly different conduction behaviors: the current for the set state shows a linear (ohmic) increase with voltage, with resistance $R = V/I$ of about 10 kΩ. On the other hand, the reset state shows a linear increase of current for very low voltage (less than about 0.2 V), followed by an almost exponential increase of current for relatively large voltage. A sudden voltage drop, or snapback, occurs at the so-called threshold voltage V_T of about 1.2 V in the figure, corresponding to a threshold current $I_T \approx 3$ μA for this technology. The voltage snapback is the result of threshold switching, that is, a characteristic negative differential resistance (NDR) effect which is generally observed in amorphous phase change materials and OTS [19, 41–50]. After threshold switching, the I-V curve of the reset state overlaps to that of the set state, most likely due to the crystallization taking place at large current and relatively long times during the DC measurements. Note that at low voltage in the ohmic regime, the reset state displays

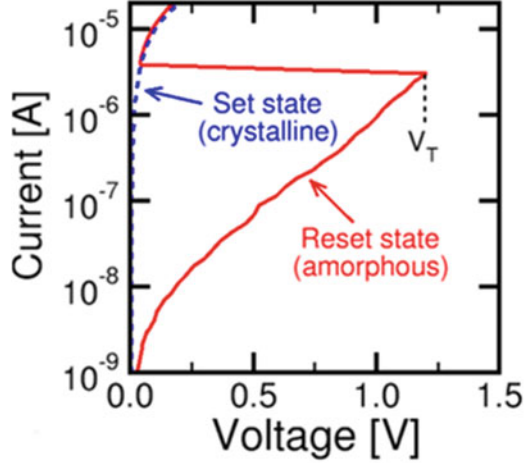


Fig. 2.7 Measured I-V curves for a PCM device in the reset state (amorphous phase) and set state (crystalline phase) [39]

a resistance of about 10 MΩ, thus resulting in a resistance ratio of about 10^3 between the two logic states in the PCM.

The linear I-V curve of the set state can be attributed to the band structure of the fcc GST, where the degenerate Fermi level causes a large concentration of free carriers contributing to drift transport. The positive activation energy of the fcc structure in Fig. 2.3 might be explained by the polycrystalline structure of the set state [51], where carriers are thermally activated in the transport over energy barriers associated to grain boundaries [52]. The variable positive activation energy in the crystalline phase was also attributed to disorder-induced localized states in the crystalline phase, according to the Anderson theory of localization [53].

2.2.2 *Poole-Frenkel Conduction in the Reset State*

The nonlinear conduction of the I-V curve of the reset state can instead be attributed to Poole-Frenkel (PF) conduction [27, 36, 37, 47, 48]. Figure 2.8 schematically illustrates the band diagram under an applied field in the amorphous GST, where carrier transport is due to thermally excited hopping between positively charged localized states [54]. In the case of relatively short distance Δz between localized states (Fig. 2.8a), the energy barrier $\Delta\phi(V)$ under an applied voltage V is given by [47]:

$$\Delta\phi(V) = \Delta\phi(0) - qF\frac{\Delta z}{2} = E_C - E_{F0} - qV\frac{\Delta z}{2u_a} \tag{2.2}$$

where $\Delta\phi(0)$ is the energy barrier under zero electric field F, E_{F0} is the equilibrium Fermi level, E_C is the maximum of the barrier equivalent to the conduction band minimum between two localized states, and u_a is the thickness of the amorphous

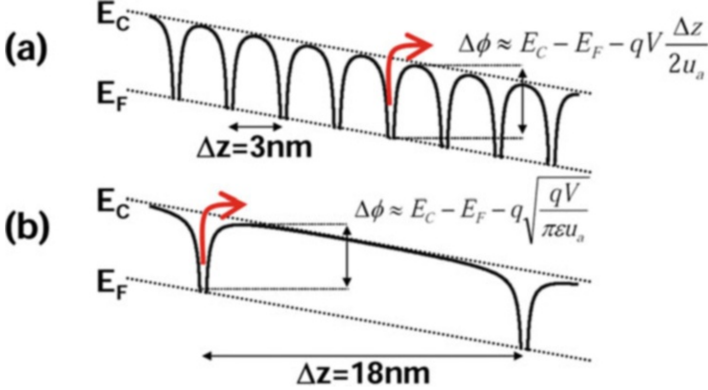


Fig. 2.8 Schematic diagram of PF conduction in amorphous GST, comparing the cases where the distance Δz between charged localized states is relatively short (**a**) or relatively long (**b**) Reprinted with premission from [27]. Copyright (2008) IEEE

volume where all the voltage V is assumed to drop. According to Eq. (2.2), the energy barrier for PF transport decreases linearly with V, which applies for relatively small Δz. As a result, the thermally activated current for relatively large V can be written as [47]:

$$I = qAN_T\frac{\Delta z}{\tau_0}e^{-\frac{E_C - E_{F0}}{kT}}e^{\frac{qV}{kT}\frac{\Delta z}{2u_a}} \tag{2.3}$$

where N_T is the concentration of localized states and T is the local device temperature. Equation (2.3) accounts for the exponential increase of I with voltage in Fig. 2.7. Note that the thermally activated forward current in the direction of the electric field must be corrected by a reverse current in the opposite direction, which is comparable to the forward contribution at small voltages. After the correction, the current is given by [47]:

$$I = 2qAN_T\frac{\Delta z}{\tau_0}e^{-\frac{E_C - E_{F0}}{kT}}\sinh\left(\frac{qV}{kT}\frac{\Delta z}{2u_a}\right) \tag{2.4}$$

where N_T is the density of localized states and T is the local device temperature (I would remove this as already explained in the previous equation). Equation (2.4) accounts for the linear behavior of current at small voltage, as well as for the exponential current at higher voltage. For relatively long Δz (Fig. 2.8b), the physical picture of conduction remains qualitatively the same, except for the voltage dependence of energy barrier which decreases with the square root of V, thus causing the current to increase exponentially with $V^{1/2}$, namely, $\log(I) \sim V^{1/2}$, in line with the traditional PF model of conduction [47]. Both regimes can be experimentally observed in PCM devices, with the conventional PF ($V^{1/2}$) regime arising preferentially for relatively small devices, whereas the V regime

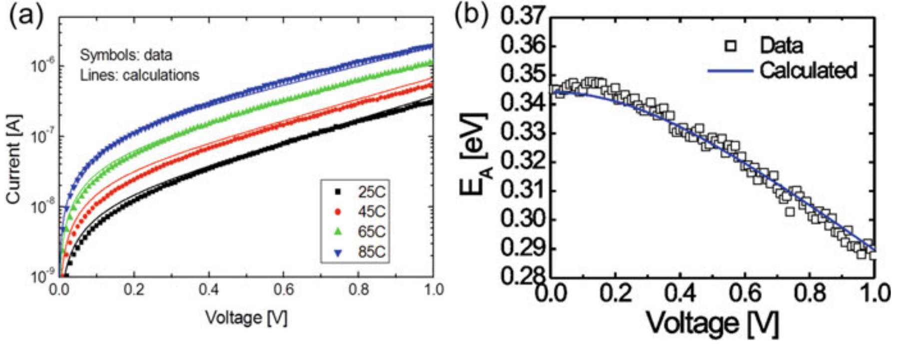


Fig. 2.9 Measured and calculated I-V curves in the reset state at variable temperature (**a**) and corresponding activation energy $E_A = d\log(I)/d(1/kT)$ as a function of V (**b**) Reprinted with permission from [47]. Copyright (2007) AIP Publishing LLC

of conduction appearing typically in larger devices [55]. A transition from V to $V^{1/2}$ regime can also be seen as a function of time from the reset process, due to the structural relaxation of the amorphous GST, also known as the drift phenomenon [54]. Initially, the I-V curve of the device shows a small Δz behavior with $\log(I) \sim V$, while at long times after reset, the conventional PF regime with $\log(I) \sim V^{1/2}$ appears. This can be attributed to the annihilation of localized states caused by the structural relaxation [54].

For a full confirmation of the physical picture of transport in Fig. 2.8, one should test the model against as many as possible experimental dependencies, such as the dependence on temperature and the impact of the programming condition. To this purpose, Fig. 2.9a shows the measured and calculated I-V curves for T increasing from 25 °C to 85 °C. There is a good agreement between the calculated I-V curves and the reported data, in both the V dependence of the individual I-V curves and the impact of temperature. As T increases, the current increases quite significantly due to the Arrhenius-type thermal activation in Eq. (2.4). These data allow to extract the activation energy at variable V, from the slope of $\log(I)$ in the Arrhenius plot [47]. Figure 2.9b shows the activation energy $E_A = d\log(I)/d(1/kT)$ as a function of V, extracted from data and calculations in Fig. 2.9a. Note that the activation energy can be viewed as the effective value of the energy barrier for carrier transport in the PF process. The decrease of E_A at increasing V is thus a clear evidence and confirmation of the PF conduction mechanism in the amorphous phase change material.

Another compelling evidence of the PF transport mechanism in the amorphous phase is obtained by testing the dependence of the subthreshold slope (STS), given by $\mathrm{STS} = d\log(I)/dV$ on the amorphous layer thickness u_a. Figure 2.10a shows the TEM pictures of the active cell region evidencing the amorphous cap in three cells programmed at increasing voltage, while Fig. 2.10b shows the corresponding I-V curves in the PF plot, where $\log(I)$ is reported as a function of $V^{1/2}$. By calculating the derivative of $\log(I)$ with respect to V in the long Δz regime, one obtains [36]:

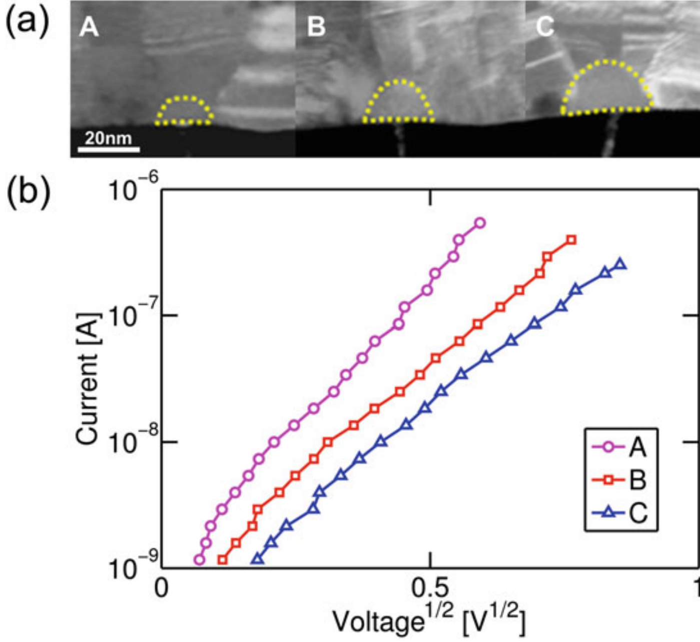


Fig. 2.10 TEM of the amorphous cap for increasing programming voltage (**a**) and corresponding measured current as a function of $V^{1/2}$ (**b**) Reprinted with permission from [36]. Copyright (2013) IEEE

$$\mathrm{STS} = \frac{q}{kT}\sqrt{\frac{q}{\pi\epsilon u_{\mathrm{a}}}} \tag{2.5}$$

where ε is the dielectric constant. Data in Fig. 2.10b indicate that the STS in the PF plot decreases with increasing u_{a}, in qualitative agreement with Eq. (2.5), thus supporting the PF model of conduction. Further support to the PF model was provided by demonstrating the inverse dependence of STS on T according to Eq. (2.5) [56].

Further analysis of the I-V curves at variable thickness allows to evidence energy landscape effects, due to the variation of energy barrier resulting from the disordered structure [36, 57, 58]. In particular, the energy landscape is at the origin of the thickness-dependent resistivity, where amorphous GST resistivity increases with thickness due to the percolative nature of PF conduction in the energy landscape. An important implication of the energy landscape effect is the nonlinear increase of resistance with thickness as the thickness of the amorphous cap increases in the PCM device [57]. This is at the basis, for instance, of the gradual increase of resistance along the programming curve, where an increase of the programming voltage leads to a gradual increase of resistance corresponding to the amorphous volume formed after quenching from the melt. The origin of this thickness-dependent resistivity is explained by Fig. 2.11, showing the calculated potential energy along the electron percolative path. The potential energy can be

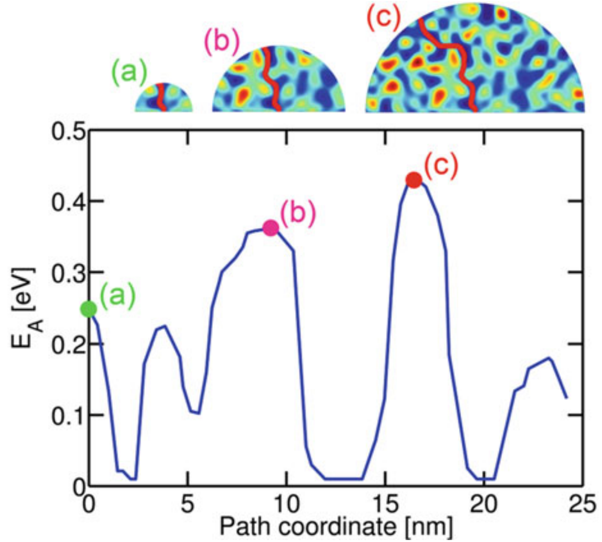


Fig. 2.11 Calculated profile of energy barrier along the percolation channel, indicating energy landscape effects where the limiting energy barrier increases with thickness. Three amorphous volumes (a, b, and c) with increasing thickness are considered. Reprinted with permission from [36]. Copyright (2013) IEEE

viewed as the energy barrier E_A which the electrons must overcome by thermal excitation in the PF transport. For a thin amorphous cap (a), the percolative path is short; thus E_A will statistically display relatively small values. As the amorphous cap is increased as in (b) or (c), there are increasing chances to cross a high E_A point in the percolative path. As a result, E_A (hence the measured resistivity) increases with the amorphous cap thickness. Energy landscape phenomena should thus be taken into account for projecting the thickness scaling of PCM.

2.3 Threshold Switching

The most distinguishing aspect of the conduction characteristics in the amorphous chalcogenide is the threshold switching, that is, the abrupt change of conductivity occurring at the threshold voltage V_T in Fig. 2.7 [41–50]. Threshold switching generally arises due to bistable S-type NDR, where two different currents can be achieved for a given voltage within a restricted voltage range. This is schematically illustrated in Fig. 2.12a showing the typical I-V curve with an OFF state at low current and an ON state at high current in correspondence of the same voltage. The bistable behavior can be understood by the fact that the same voltage drop across the same semiconductor thickness can be sustained with two different band diagrams: in the OFF state (Fig. 2.12b), the voltage drops homogeneously across the device thickness, resulting in a relatively low electric field inducing equilibrium carrier transport, such as the PF conduction process in Fig. 2.8. On the other hand, the ON state (Fig. 2.12c) consists of a non-equilibrium configuration of bands with heavily nonhomogeneous distribution of the electric field. The field is extremely high within an initial dead layer, where the high field induces injection of carriers in

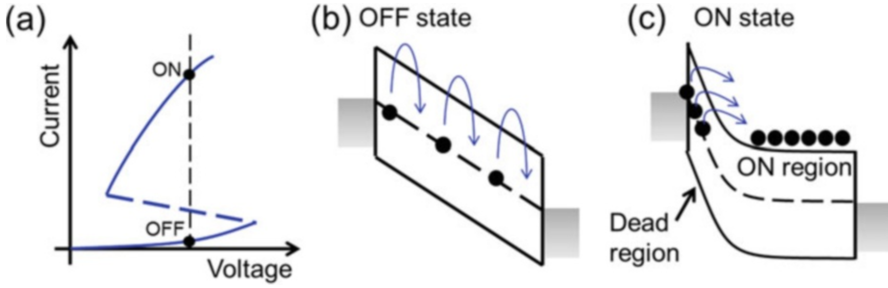


Fig. 2.12 Schematic illustration of NDR mechanism at the basis of threshold switching, showing the typical I-V characteristics for NDR (**a**) and the band diagrams for the OFF state (**b**) and the ON state (**c**)

the following ON region, where the field is instead relatively low, thanks to the large concentration of injected carriers. Alternatively, injection might lead to transfer of carriers to higher energy, thus enhancing their mobility. Current continuity in the ON and dead layers is ensured since the current is proportional to the product of carrier density, carrier mobility, and electric field; thus a high field and low concentration/mobility in the dead layer are balanced by the low field and high concentration/mobility in the ON layer.

The injection mechanisms in the dead layer of Fig. 2.12c can arise from different mechanisms, such as tunneling [47] or even PF conduction itself [48], as a result of the carrier heating phenomenon illustrated in Fig. 2.13. In fact, the PF mechanism involves the thermal excitation of carriers over an energy barrier for transport. At low field (Fig. 2.13a), carriers can relax their energy after a hopping event; thus, the energy distribution of carriers remains that of the equilibrium state, where roughly all states are filled below E_F while roughly all states are empty above E_F. One should recall that a large concentration of localized states exists in the amorphous semiconductor band gap and sustains the PF transport process. At sufficiently high electric fields (Fig. 2.13b), there is no sufficient time for the carrier energy to relax between the capture and emission events [59–61]; thus, carriers can gain a significant amount of energy resulting in a hot (non-equilibrium) electron distribution. This is similar to the hot electron effects arising in a FET channel at high electric fields [62], except that the non-equilibrium distribution mostly affects the density of carriers in localized states, rather than in the band of extended states. Due to the exponential energy dependence of electron mobility in the PF transport, the conductivity can be significantly enhanced, thus resulting in a reduced voltage drop in the region where carriers have high energy, in agreement with the simplified conceptual picture of Fig. 2.12c. Note that this picture only assumes PF at low/high field as the origin of threshold switching and thus is capable of explaining threshold switching in a wide class of materials, e.g., phase change chalcogenides, OTS materials, and in principle any other material sharing the PF mechanism as the dominant conduction process [48].

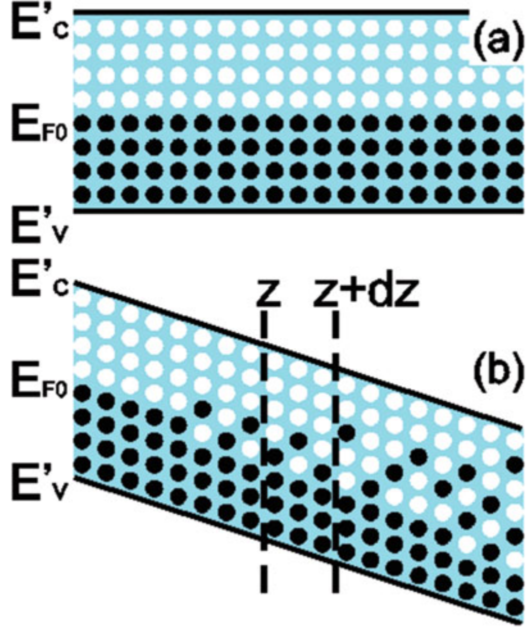


Fig. 2.13 Schematic band diagram illustrating the mechanism of energy gain describing threshold switching. At zero voltage, electrons occupy all states below the equilibrium of Fermi level E_{F0} (**a**), while the application of a relatively high voltage causes the gain of energy of carriers due to (**b**) Reprinted with permission from [48]. Copyright (2008) American Physical Society

Based on the physical picture in Fig. 2.13, a numerical model was developed to mathematically reproduce the NDR and consequent threshold switching [48]. According to Fig. 2.13, the model relies on an energy balance for carriers (electrons in the figure) in a layer of infinitesimal thickness dz. The energy flow input into the infinitesimal layer can be written as $J(z)E_F(z)/q$, namely, the carrier flow multiplied by the average carrier flow equal to the local quasi Fermi energy E_F. The latter can be different from the equilibrium value E_{F0} as a result of the energy gain induced by the field. Within the layer, carriers receive a contribution to the energy equal to the electrostatic energy loss, namely, $dE_{gain} = qFdz$. On the other hand, scattering with phonons and defects leads to an energy loss given by $dE_{loss} = (E_F\text{-}E_{F0})n_T dz/\tau_{rel}$, where n_T is the effective density of trapped carrier contributing to the energy flow and τ_{rel} is a characteristic relaxation time constant describing the carrier tendency to thermalize their distribution from non-equilibrium [48]. By combining the four contributions (energy input, energy output, energy increase, and energy loss in the layer), one obtains a differential equation for the carrier energy, given by:

$$\frac{dE'_F}{dz} = qF - \frac{qn_T}{J\tau_0}\frac{E'_F - E_F}{\tau_{rel}} \tag{2.6}$$

which allows to calculate the E_F profile for any applied current density J [48]. Equation (2.6) can be solved by explicating the relationship between local field F and current density, e.g., based on the PF transport equation, under the continuity condition that J remains uniform over the entire thickness u_a.

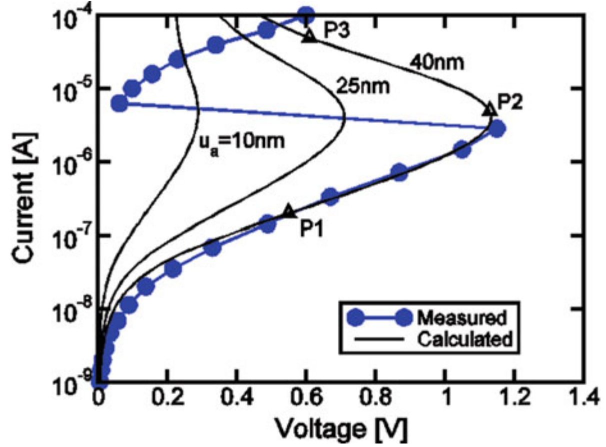


Fig. 2.14 Measured and calculated I-V curves for a PCM device, showing threshold switching induced by S-type NDR. Calculations were performed for variable thickness of the amorphous layer. Reprinted with permission from [48]. Copyright (2008) Americal Physical Society

Figure 2.14 shows the I-V curves obtained from measurements and calculations using Eq. (2.6) for various values of the amorphous GST thickness u_a. Calculations agree with the measured data for a thickness $u_a = 40$ nm, which agrees with the typical thickness of the amorphous cap in a PCM device after full reset operation. Calculations show the characteristic S-type NDR behavior as a result of the increase of the carrier energy E'_F as described in Fig. 2.13. Three points are marked in Fig. 2.14, corresponding to subthreshold condition (P1), where the field is too low to induce any significant energy increase of carriers; thus, $E_F = E_{F0}$ over the whole thickness. At the threshold point (P2), the energy gain is sufficiently high to induce a significant boost of conductivity in the ON region. Starting from this condition, any further increase of current density results in an increased field in the dead layer, a higher energy gain in the ON region, and a correspondingly smaller field in the ON region. As a result, the voltage decreases at increasing current, which is at the basis of the NDR behavior in Fig. 2.14. Point P3 at high current in the NDR regime is an example of this situation, which can be visualized by Fig. 2.12c, while P1 at almost equal voltage drop can be represented by Fig. 2.12b. The threshold point P2 thus marks the conditions (current density and voltage V_T) for which the energy gain of the electron population reaches a critical value, around kT, which enables sufficient enhancement of conductivity to trigger the NDR. A detailed analysis shows that a fixed energy gain is equivalent to a constant power density $P''' = FJ$ within the device. This condition can be used for a fast prediction of threshold point in PCM devices.

Note again the large generality of this description which results from the energy balance of Eq. (2.6) without invoking any additional mechanisms except PF transport under non-equilibrium distribution of energy induced by the high electric field. Note also that the contribution of additional mechanisms, such as Joule heating [38], impact ionization [19, 45], polaron instability [46], etc., may not be ruled out from this analysis. A different view is proposed in Chap. 4 where carrier

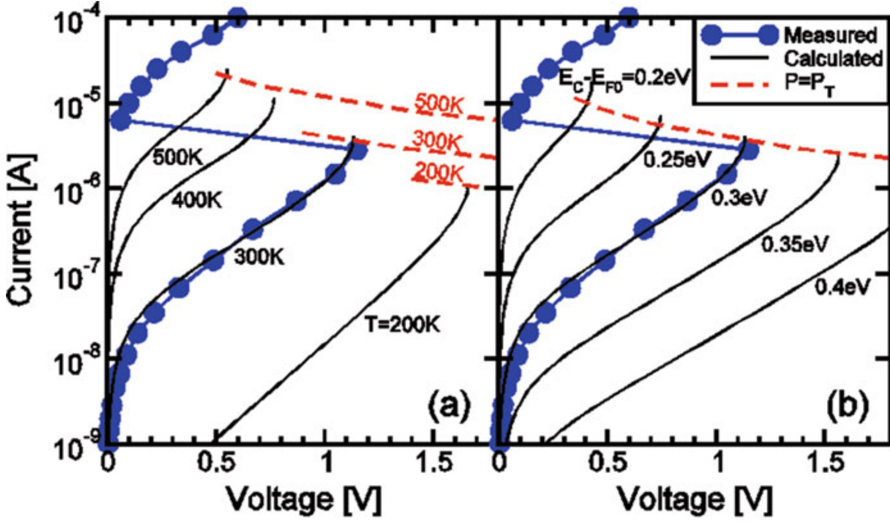


Fig. 2.15 Measured and calculated I-V curves for a PCM device for variable temperature (**a**) and variable E_C-E_{F0} (**b**) Lines at constant power density are also shown for reference of the constant-power rule of threshold switching. Reprinted with permission from [48]. Copyright (2008) American Physical Society

generation-recombination balance is at the basis of the switching event, which can be conveniently implemented in a finite element-based device simulator tool.

Figure 2.15a shows the calculated I-V curves for variable temperature (a). A reference experimental characteristic measured at 300 K for GST is also shown for reference. The NDR region of the I-V curves was deleted for clarity, and the I-V curves were shown only up to the threshold voltage V_T. Calculations show that the threshold voltage V_T decreases with increasing T, as a result of the Arrhenius-type increase of subthreshold PF leakage. In fact, to maintain a constant power density for switching, V_T decrease as the subthreshold current increases. This results in an approximately linear decrease of V_T with T.

Figure 2.15b shows the calculated I-V curves for variable E_C-E_{F0}, which corresponds to about half of the band gap assuming that the Fermi level is located at midgap. The PF current decreases at increasing E_C-E_{F0}, thus causing V_T to increase as a result of the constant-power rule. This is generally demonstrated experimentally, as phase change materials with a low leakage current display a relatively large activation energy for conduction, as well as a relatively large V_T. This provides relevant design guidelines to engineer the phase change material to match a certain circuit requirement for V_T.

2.4 Switching Dynamics

While most modeling and characterization studies aim at describing the DC behavior of switching, namely, the threshold voltage and current under quasi-static DC-type measurement conditions, time-dependent analysis of switching has been attempted less frequently. The time-dependent switching dynamics can be studied with reference to two fundamental parameters: the delay time τ_D, namely, the time needed to initiate the switching process since the first application of the voltage, and the switching time, τ_S, namely, the time for the current to rise from the subthreshold level to the ON state.

The switching time can be viewed as the time to switch from the OFF-state (uniform field) configuration in Fig. 2.12b to the ON-state configuration (nonuniform field) in Fig. 2.12c, requiring electron heating and establishment of the high-energy/high-concentration conditions in the ON region. This evolution is governed by a positive feedback effect, in that a momentary increase of carrier energy results in a higher conductivity in the ON region, thus causing a higher current and a corresponding increase of the voltage drop across the dead region to sustain the increased current. A higher electric field across the dead region results in (i) an increase of the energy in the OFF region, causing a further increase of current, and (ii) a decrease of the voltage drop across the ON region to maintain the constant voltage across the PCM device. Ideally, such a positive feedback loop would lead to an indefinite increase of current causing final destructive failure of the device, but in practice the series resistance in the PCM device, e.g., caused by the confined bottom electrode, would limit the maximum current, thus allowing to reach a final steady state as sketched in Fig. 2.12c.

Attempts to measure the switching time were not successful to date, possibly due to the positive feedback loop and the correspondingly very fast switching transient. Calculations of the switching time were carried out by using an analytical model of threshold switching based on tunneling from gap states to the conduction band [47]. Figure 2.16 shows the calculated I-V curves with the transient trajectory during time-dependent switching (a) and the calculated current (top), electric field (center), and carrier concentration (bottom) as a function of time [63]. The I-V trajectory during switching assumes a load resistance $R_L = 3\ \text{k}\Omega$ in series and includes three representative points, corresponding to initial equilibrium conditions in the OFF state (i), transient state during the switching event (ii), and final ON state at high current (iii). At time zero, the device is biased with an applied voltage in the OFF state (i) with subthreshold current flowing and a negligible concentration of excess carriers δn_{T2} in the ON region. Equal electric fields are present in the dead and ON regions due to the equilibrium conditions. However, the OFF-state field injects a gradually increasing concentration of excess carriers in the ON region as apparent from the transient of δn_{T2} in Fig. 2.16b (bottom). As δn_{T2} becomes comparable to the pre-existing (equilibrium) carrier concentration, the current starts to increase, resulting in the switching event. As the current increases during switching (ii), the electric field in the dead region F_{OFF} increases, while the electric

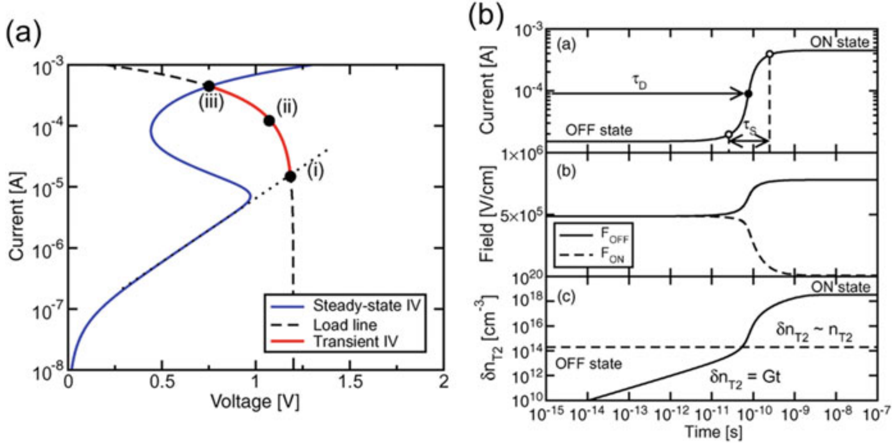


Fig. 2.16 Calculated I-V curves and trajectory during the switching transient (**a**) and corresponding current (top), electric field (center), and carrier concentration (bottom) as a function of time (**b**) Note the gradual increase of excess carriers δn_{T2} which triggers the switching event with the corresponding steep rise of current δn_{T2} and discrimination of electric fields in the dead and ON regions. Reprinted with permission from [63]. Copyright (2010) IEEE

field in the ON region F_{ON} decreases, thus establishing the nonuniform field distribution in Fig. 2.12c. Steady state is eventually reached in the ON state (iii).

The simulated current transient in Fig. 2.16b (top) allows to discriminate between delay and switching transients: the delay time τ_D can be defined as the time for current to reach half of the switching swing, while the switching time τ_S can be defined as the net time for the current to switch from the OFF to the ON state, represented by 10% and 90% of the whole current swing. Physically, the delay transient in the simulation is described by the time to build up a concentration of extra carriers δn_{T2} comparable to the equilibrium concentration, while the switching time is the time for the positive feedback-driven steep increase of the current. Note that τ_D and τ_S are both extremely short, thus posing fundamental difficulties in time-resolved measurements of these times.

More detailed analysis indicates a key role of the parasitic capacitance which is inevitably present in any PCM device during measurements [63]. In general, a parasitic capacitance C of few pF may be expected to affect a conventional experimental setup for electrical measurements [64]. The parasitic capacitance affects the measurements in two ways: (i) the delay time is increased by the RC delay in the circuit, resulting in longer τ_D given approximately by the time to build up a charge $Q_i = CV_A$ within the capacitance, and (ii) the current increases in excess of what is expected from the load-line trajectory in Fig. 2.16, due to the parasitic overshoot [63, 64]. The latter phenomenon is due to the fact that the time to discharge the parasitic capacitance C is approximately given by $R_L C$, namely, the RC time of the circuit. Since this is generally much longer than τ_S, the voltage across the PCM remains constant during switching, instead of adjusting to the lower

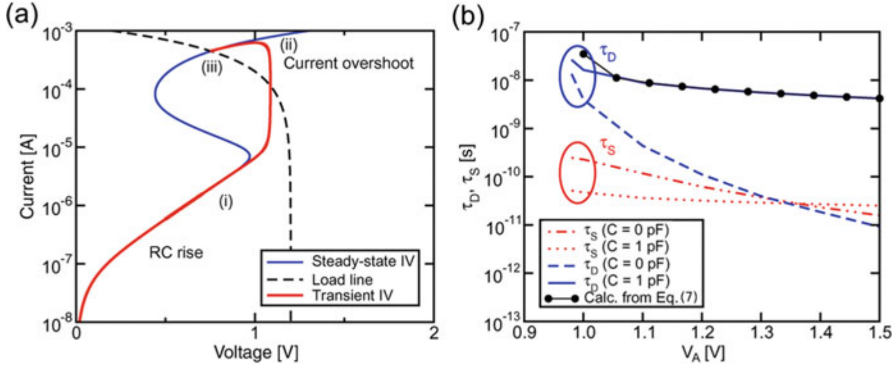


Fig. 2.17 Calculated I-V curves and trajectory during the switching transient for nonzero parasitic capacitance (**a**) and calculated τ_D and τ_S as a function of applied voltage (**b**) comparing situations with and without parasitic capacitance. Reprinted with permission from [63]. Copyright (2010) IEEE

value according to the load line in Fig. 2.16a. As a result of the larger voltage across the PCM, switching time is shorter than the calculated τ_S in Fig. 2.16b, which assumed a decreasing voltage during switching. Also, the maximum current reached after switching is larger than the case of zero capacitance, which accounts for the term "overshoot."

Figure 2.17 shows the calculated I-V curve and the switching trajectory for a PCM device with a parasitic capacitance $C = 1$ pF (a) and the corresponding τ_D and τ_S as a function of applied voltage with and without a parasitic capacitance $C = 1$ pF (b). The switching characteristics show a steep transition with constant voltage taking place in a time τ_S, followed by a readjustment of voltage taking place in a longer time R_LC. Due to the first stage of vertical switching, the current shows an overshoot effect. The parasitic capacitance results in a longer delay time which is dominated by the RC delay according to the formula:

$$\tau_D = R_L C \log \frac{1}{1 - V_T/V_A} \tag{2.7}$$

which is calculated as the time to reach close to V_T in a R_LC circuit under an applied voltage V_A [63]. Calculation from Eq. (2.7) is also shown in the figure for comparison. The switching time instead is decreased by the parasitic capacitance due to the constant voltage during the switching transient. Note however that the switching time is always in the sub-nanosecond range, which makes it difficult to attempt an experimental evaluation of the switching time.

More detailed studies of the delay time indicate that there are two regimes, namely, RC-limited delay time at relatively high voltage, as shown in Fig. 2.17 and described by Eq. (2.7), and the fluctuation-limited delay time at relatively low voltage [65]. Figure 2.18a shows the measured voltage across the PCM during switching at relatively low voltage close to V_T [66]. The sharp decrease of voltage in the figure marks threshold switching, as the PCM resistance decreases, thus

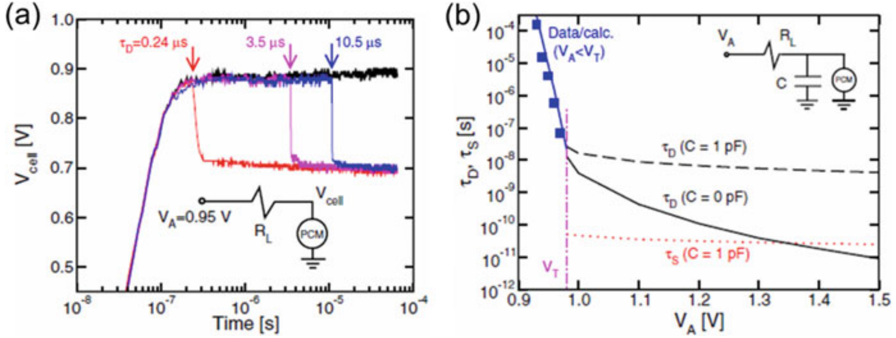


Fig. 2.18 Measured voltage V_{cell} across the PCM device (see inset) as a function of time indicating stochastic switching with variable delay time τ_{D} (**a**) and average τ_{D} as a function of applied voltage V_{A}, compared with calculated τ_{D} and τ_{S} at various parasitic capacitance C (**b**) Reprinted with permission from [66]. Copyright (2010) IEEE

changing the voltage drop in the voltage divider circuit shown in the inset. In these types of measurements, threshold switching occurs at relatively long times, i.e., in the few µs time range, compared to the few ns range in Fig. 2.17, where the shorter delay was due to the higher voltage overdrive across the device. Three repeated measurements are shown in Fig. 2.18a, indicating a broad distribution of τ_{D} due to stochastic switching. A detailed analysis of the distribution of τ_{D} at variable voltage indicates a Weibull distribution with relatively low shape factor, i.e., the slope of the distribution on the Weibull plot [65]. This behavior can be understood by current fluctuations obeying 1/f power spectral density (PSD), where a "lucky" fluctuation above a certain threshold current I_{T} triggers threshold switching. Simulations by 1/f current noise models support this picture in terms of both the statistical distribution and the average value of τ_{D} as a function of the applied voltage V_{A}, as shown in Fig. 2.18b [66]. Data and calculations indicate an extremely steep slope of about 13 mV/dec. This type of analysis allows for a nonambiguous definition of the threshold voltage V_{T}, as the voltage marking the separation of the fluctuation-limited regime of τ_{D} and the high-voltage regime which is controlled by either RC delay or by the buildup of the non-equilibrium distribution of carriers in the PCM material. The accurate understanding of delay and switching phenomena in Fig. 2.18 allows for a careful design of PCM in several conditions (read, write, disturb) and applications, such as memory [8–10], logic [67–69], and non-Boolean computation for future data-intensive computing memories [70–72].

2.5 Time-Dependent Transport

While threshold switching describes the time-dependent transition from high resistance to low resistance due to the bistable nature of the conduction characteristic, resistance drift is a physical process responsible for the transition to high resistance. This is shown in Fig. 2.19a reporting the resistance for a PCM in the amorphous phase as a function of time measured with respect to the reset pulse [27]. The resistance increases according to a power law, namely [73–75]:

$$R = R_0 \left(\frac{t}{t_0} \right)^{\nu} \tag{2.8}$$

where R_0 and t_0 are constant and the exponent ν is about 0.1. The set-state resistance is also shown in the figure, indicating a stable resistance with negligible drift. Resistance drift in the amorphous phase can be explained by structural relaxation taking place in the non-equilibrium disorder phase, as illustrated in Fig. 2.19b. Here, defects remained from the liquid phase after quenching represent high-energy configurations, which the system tends to anneal by structural relaxation. For instance, a dangling bond relaxes by fixing itself to an atom to complete a chain. Defects which might play a critical role in the structural relaxation of amorphous GST are wrong bonds, such as Ge-Ge bonds and chains which are largely unstable [76]. Structural relaxation is universally observed in disordered materials, such as amorphous Si [77, 78], amorphous Ge [79], amorphous SiC [80], and amorphous hydrogenated carbon [81]. Relaxation should not be viewed as a phase transition, since the structure remains amorphous and disordered, while the resistance and band gap increase [74, 75]. In fact, in contrast to crystallization, the energy barrier E_A for structural relaxation in Fig. 2.19b belongs to a wide distribution, from relatively small energies to large energies close to the crystallization energy barrier, e.g., about 2.6 eV for GST. As a result, structural relaxation is seen even at room

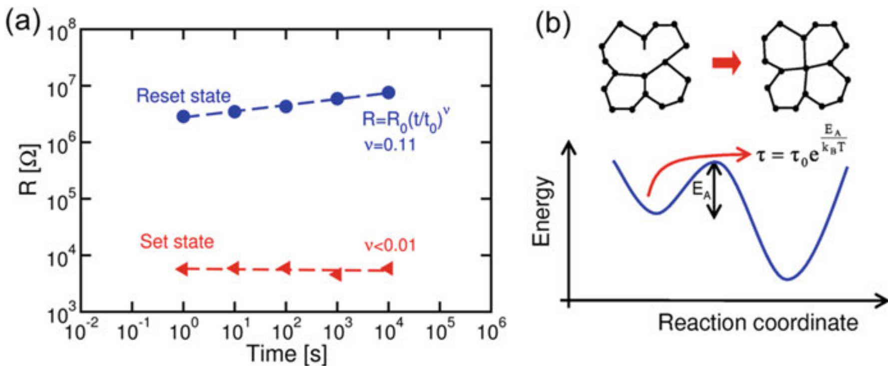


Fig. 2.19 Measured resistance R as a function of time after programming (**a**) and schematic illustration of structural relaxation (**b**) The power law drift of resistance for the reset state is due to structural relaxation, where defects are annealed by thermally accelerated relaxation to lower energy states. Reprinted with permission from [73]. Copyright (2007) IEEE

temperature, as shown by the resistance drift in Fig. 2.19a, due to the relatively small energy barriers in the distribution of E_A. The broad distribution of energy barriers is at the origin of the power law time dependence of resistance drift, based on the Arrhenius relationship between transition time and E_A [82].

Structural relaxation results in a consolidation of bonding in the amorphous state, which is reflected by an increasing band gap [40, 54] and a correspondingly decreasing dielectric constant [27]. The increased band gap leads to a higher energy barrier for PF conduction, thus causing an increase of resistance and a decrease of leakage current. The lower leakage current also causes a higher threshold voltage, which shows a drift behavior where V_T increases linearly with the logarithm of time [55, 75]. In general, the increase of the reset state resistance does not raise any reliability issue, as the resistance window between set and reset states increases with time. However, the increases V_T may eventually make set-state programming difficult, if V_T increases above the nominal voltage for set process. This possible problem is easily fixed by adjusting the programming voltage to overcome possible V_T drift effects. The main impact of drift on PCM reliability is seen in the case of multilevel storage, where the PCM is programmed at more than two levels to enable the storage of multiple bits per cell. For instance, four resistance levels, corresponding to set, reset, and two intermediate resistance levels, allow the representation of two bits in an individual physical cell. Due to the presence of a mixed phase in the intermediate states, resistance displays a measurable drift [83], thus causing broadening of the programmed distributions with time which makes multilevel storage difficult in PCM. In some PCM materials, such as Ge-rich GST which is designed to achieve a higher crystallization temperature for embedded memory and automotive applications, the set state also displays a significant drift due to the presence of substantial disordered phase after set transition [31, 52]. Set-state drift may result in a resistance window closure that may negatively affect reliability.

To overcome the resistance drift problem in PCM, some solutions were proposed at the level of both read/program algorithms and cell architecture. Figure 2.20 shows the measured R as a function of time for a PCM in the reset state annealed at room temperature or at 120 °C. The resistance was measured at a given read current I_{read}, instead of a fixed read voltage, and measurement results are shown for increasing I_{read}. As I_{read} increases, the resistance value decreases, as a result of the exponential-like increase of the PF current. Most importantly, resistance drift decreases significantly for increasing I_{read}, which can be again understood by the exponential-like increase of the current. In fact, assuming that the current is given by $I = I_0 \exp(V_{read}/V_0)$, and assuming that drift is simply described by a decrease of I_0, the relative resistance increases between time t_1 and time t_2 and is given by [54]:

$$R(t_2)/R(t_1) = \log(I_{read}/I_0(t_2))/\log(I_{read}/I_0(t_1)),$$

which tends to $R(t_2)/R(t_1) = 1$ for $I_{read} >> I_0(t_1) > I_0(t_2)$, thus resulting in a strong attenuation of the resistance drift with respect to the case of small I_{read}, of the case of read at a constant voltage, where $R(t_2)/R(t_1) = I_0(t_1)/I_0(t_2)$ [54]. Similar algorithms based on alternative read metrics were proposed to minimize drift and allow multilevel cell storage in PCM [84].

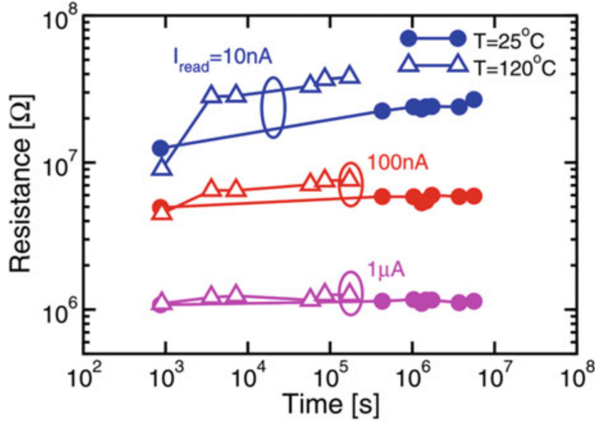


Fig. 2.20 Measured resistance R as a function of time at annealing temperature equal to T = 25°C and T = 120°C. The resistance was measured at room temperature and increasing current, namely I_{read} = 10 nA, 100 nA, and 1 μA. As I_{read} increases, the drift slope decreases thanks to the exponential increase of the conduction characteristics. Reprinted with permission from [27]. Copyright (2008) IEEE

Drift can also be mitigated by adopting alternative programming mechanisms, such as the bipolar switching. In bipolar switching, different resistance states are obtained by inducing ionic migration, instead of a real phase change [39]. In particular, the high resistance state is obtained by applying a high negative voltage at the top electrode, resulting in a depletion of cation species (Ge, Sb) at the bottom electrode contact. As the high resistance does not rely on an amorphous region, structural relaxation and drift are fully suppressed in the reset state [39]. Alternative cell architectures have also been proposed to solve the drift problem. For instance, a core-shell structure of the PCM, where a phase change material core pillar is surrounded by a metallic surfactant layer shell, was shown to display less drift, as a result of the parallel current path in the surfactant layer [85]. In fact, the current can sneak around the amorphous region through the surfactant layer, thus preventing any impact of structural relaxation to resistance. The drift coefficient could thus be significantly reduced, although the resistance window also decreases as a result of the surfactant conduction path with relatively low resistance. Combining material, device, and algorithm solutions might enable reliable multilevel storage for ultrahigh-density PCM in the future.

2.6 Summary

PCM operation during read and write relies on a careful engineering of the interplay between electrical, thermal, and structural phenomena at the nanoscale. In this interplay, subthreshold conduction, threshold switching, and ON/OFF state transport properties dictate most of the PCM working conditions.

This chapter has reviewed the most significant PCM electrical properties affecting resistance window, write and read voltage values, and dynamic response. It is shown that the resistance window relies on the band structure of the phase change semiconductor material in the crystalline and amorphous phase. The transport properties of these two phases have been described with reference to the carrier drift or hopping conduction mechanisms. Threshold switching has been reviewed in deep detail, covering its fundamental mechanism, its steady-state description, and its dynamics, supporting the fundamental role of a detailed electrical characterization for accurate modeling and design of PCM memory circuits.

Acknowledgments This work was supported in part by the European Research Council Consolidator Grant ERC-2014-CoG-648635-RESCUE.

References

1. M. Wuttig, N. Yamada, Phase-change materials for rewriteable data storage. Nat. Mater. **6**, 824–832 (2007)
2. H.-S.P. Wong, S. Raoux, S. Kim, J. Liang, J.P. Reifenberg, B. Rajendran, M. Asheghi, K.E. Goodson, Phase change memory. Proc. IEEE **98**, 2201 (2010)
3. S. Raoux, W. Welnic, D. Ielmini, Phase change materials and their application to non-volatile memories. Chem. Rev. **110**, 240–267 (2010)
4. D. Ielmini, A.L. Lacaita, Phase change materials in non-volatile storage. Mater. Today **14**, 600–607 (2011)
5. S. Raoux, D. Ielmini, M. Wuttig, I.V. Karpov, Phase change materials. MRS Bull. **37**, 118–123 (2012)
6. N. Yamada, E. Ohno, K. Nishiuchi, N. Akahira, M. Takao, J. Appl. Phys. **69**, 2849 (1991)
7. D. Mantegazza, D. Ielmini, A. Pirovano, A.L. Lacaita, E. Varesi, F. Pellizzer, R. Bez, Explanation of programming distributions in phase-change memory arrays based on crystallization time statistics. Solid State Electron. **52**, 584–590 (2008)
8. F. Pellizzer, A. Pirovano, F. Ottogalli, M. Magistretti, M. Scaravaggi, P. Zuliani, M. Tosi, A. Benvenuti, P. Besana, S. Cadeo, T. Marangon, R. Morandi, R. Piva, A. Spandre, R. Zonca, A. Modelli, E. Varesi, T. Lowrey, A. Lacaita, G. Casagrande, P. Cappelletti, R. Bez, Symp. VLSI. Tech. Dig., 18 (2004)
9. F. Pellizzer, A. Benvenuti, B. Gleixner, Y. Kim, B. Johnson, M. Magistretti, T. Marangon, A. Pirovano, R. Bez, G. Atwood, Symp. VLSI Tech. Dig, 122 (2006)
10. G. Servalli, IEDM Tech. Dig, 113 (2009)
11. B.-S. Lee, J.R. Abelson, S.G. Bishop, D.-H. Kang, B.-K. Cheong, K.-B. Kim, Investigation of the optical and electronic properties of $Ge_2Sb_2Te_5$ phase change material in its amorphous, cubic, and hexagonal phases. J. Appl. Phys. **97**, 093509 (2005)
12. T. Ide, M. Suzuki, M. Okada, Jpn. J. Appl. Phys. **34**(Part 2), L529 (1995)
13. D.V. Tsu, J. Vac. Sci. Technol. A **17**, 1854 (1999)
14. S.Y. Kim, S.J. Kim, H. Seo, M.R. Kim, Jpn. J. Appl. Phys. **37**(Part 138), 1713 (1999)
15. E. García-García, A. Mendoza-Galván, Y. Vorobiev, E. Morales-Sánchez, J. González-Hernández, G. Martínez, B.S. Chao, J. Vac. Sci. Technol. A **17**, 1805 (1999)
16. R. Liang, C. Peng, K. Nagata, K. Daly-Flynn, M. Mansuripur, Appl. Opt. **41**, 370 (2002)
17. H.B. Yao, L.P. Shi, T.C. Chong, P.K. Tan, X.S. Miao, Jpn. J. Appl. Phys. **42**(Part 1), 828 (2003)

18. S.H. An, X. Li, S.Y. Kim, Jpn. J. Appl. Phys. **43**(Part 1), 1006 (2004)
19. A. Pirovano, A.L. Lacaita, A. Benvenuti, F. Pellizzer, R. Bez, IEEE Trans. Electron Devices **51**, 452 (2004)
20. N.F. Mott, E.A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd edn. (Clarendon, Oxford, New York, 1979)
21. J. Tauc, R. Grigorovici, A. Vancu, Phys. Status Solidi **15**, 627 (1966)
22. T. Kato, K. Tanaka, Electronic properties of amorphous and crystalline $Ge_2Sb_2Te_5$ films. Jpn. J. Appl. Phys. **44**, 7340–7344 (2005)
23. J.-J. Kim, K. Kobayashi, E. Ikenaga, M. Kobata, S. Ueda, T. Matsunaga, K. Kifune, R. Kojima, N. Yamada, Electronic structure of amorphous and crystalline (GeTe)1−x(Sb2Te3)x investigated using hard x-ray photoemission spectroscopy. Phys. Rev. B **76**, 115124 (2007)
24. N. Yamada, T. Matsunaga, Structure of laser-crystallized Ge Sb Te sputtered thin films for use in optical memory. J. Appl. Phys. **88**, 7020–7028 (2000)
25. T.H. Lee, S.R. Elliott, Structural role of vacancies in the phase transition of Ge2Sb2Te5 memory materials. Phys. Rev. B **84**, 094124 (2011)
26. L.E. Shelimova, O.G. Karpinskii, P.P. Konstantinov, M.A. Kretova, E.S. Avilov, V.S. Zemskov, Inorg. Mater. **37**, 342 (2001)
27. D. Ielmini, D. Sharma, S. Lavizzari, A.L. Lacaita, Physical mechanism and temperature acceleration of relaxation effects in phase change memory cells. IEEE Int. Reliab. Phys. Symp. (IRPS), 597–603 (2008)
28. M.H. Cohen, H. Fritzsche, S.R. Ovshinsky, Simple band model for amorphous semiconducting alloys. Phys. Rev. Lett. **22**, 1065–1068 (1969)
29. R.A. Street, N.F. Mott, States in the gap in glassy semiconductors. Phys. Rev. Lett. **35**, 1293 (1975)
30. S.R. Ovshinsky, Localized states in the gap of amorphous semiconductors. Phys. Rev. Lett. **36**, 1469–1472 (1976)
31. N. Ciocchini, E. Palumbo, M. Borghi, P. Zuliani, R. Annunziata, D. Ielmini, Unified reliability modeling of Ge-rich phase change memory for embedded applications. IEDM Tech. Dig, 581 (2013)
32. Y. Sasago, M. Kinoshita, T. Morikawa, K. Kurotsuchi, S. Hanzawa, T. Mine, A. Shima, Y. Fujisaki, H. Kume, H. Moriya, N. Takaura, K. Torii, Symp. VLSI Tech. Dig., 24 (2009)
33. D. Kau, S. Tang, I.V. Karpov, R. Dodge, B. Klehn, J.A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee, T. Langtry, K.-W. Chang, C. Papagianni, J. Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro, G. Spadini, IEDM Tech. Dig., 617 (2009)
34. G.W. Burr, K. Virwani, R.S. Shenoy, A. Padilla, M. BrightSky, E.A. Joseph, M. Lofaro, A.J. Kellock, R.S. King, K. Nguyen, A.N. Bowers, M. Jurich, C.T. Rettner, B. Jackson, D.S. Bethune, R.M. Shelby, T. Topuria, N. Arellano, P.M. Rice, B.N. Kurdi, K. Gopalakrishnan, Large-scale (512kbit) integration of multilayer-ready access-devices based on mixed-ionic-electronic-conduction (MIEC) at 100% yield. Symp. VLSI Tech. Dig., 41–42 (2012)
35. A. Calderoni, M. Ferro, E. Varesi, P. Fantini, M. Rizzi, D. Ielmini, Understanding over-reset transition in phase change memory (PCM) characteristics. IEEE Electron Device Lett. **33**, 1267–1269 (2012)
36. M. Rizzi, M. Ferro, P. Fantini, D. Ielmini, Energy landscape model of conduction and phase transition in phase change memories. IEEE Trans. Electron Devices **60**, 3618–3624 (2013)
37. Y.H. Shih, M.H. Lee, M. Breitwisch, R. Cheek, J.Y. Wu, B. Rajendran, Y. Zhu, E.K. Lai, C.F. Chen, H.Y. Cheng, A. Schrott, E. Joseph, R. Dasaka, S. Raoux, H.L. Lung, C. Lam, Understanding amorphous states of phase-change memory using Frenkel-Poole model. IEDM Tech. Dig., 753–756 (2009)
38. M. Le Gallo, A. Athmanathan, D. Krebs, A. Sebastian, Evidence for thermally assisted threshold switching behavior in nanoscale phase-change memory cells. J. Appl. Phys. **119**, 025704 (2016)

39. N. Ciocchini, M. Laudato, M. Boniardi, E. Varesi, P. Fantini, A.L. Lacaita, D. Ielmini, Bipolar switching in chalcogenide phase change memory. Sci. Rep. **6**, 29162 (2016)
40. P. Fantini, S. Brazzelli, E. Cazzini, A. Mani, Band gap widening with time induced by structural relaxation in amorphous $Ge_2Sb_2Te_5$ films. Appl. Phys. Lett. **100**, 013505 (2012)
41. S.R. Ovshinsky, Reversible electrical switching phenomena in disordered structures. Phys. Rev. Lett. **21**, 1450 (1968)
42. H. Fritzsche, Physics of instabilities in amorphous semiconductors. IBM J. Res. Dev. **13**, 515–521 (1969)
43. A.E. Owen, J.M. Robertson, Electronic conduction and switching in chalcogenide glasses. IEEE Trans. Electron Devices **20**, 105–122 (1971)
44. A.C. Warren, Reversible thermal breakdown as a switching mechanism in chalcogenide glasses. IEEE Trans. Electron Devices **20**, 123–130 (1973)
45. D. Adler, M.S. Shur, M. Silver, S.R. Ovshinsky, Threshold switching in chalcogenide-glass thin films. J. Appl. Phys. **51**, 3289–3309 (1980)
46. D. Emin, Current-driven threshold switching of a small polaron semiconductor to a metastable conductor. Phys. Rev. B **74**, 035206 (2006)
47. D. Ielmini, Y. Zhang, Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices. J. Appl. Phys. **102**, 054517 (2007)
48. D. Ielmini, Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses. Phys. Rev. B **78**, 035308 (2008)
49. V.G. Karpov, Y.A. Kryukov, I.V. Karpov, M. Mitra, Field-induced nucleation in phase change memory. Phys. Rev. B **78**, 052201 (2008)
50. S. Hudgens, Progress in understanding the Ovshinsky effect: Threshold switching in chalcogenide amorphous semiconductors. Phys. Status Solidi B **249**, 1951–1955 (2012)
51. G.W. Burr, P. Tchoulfian, T. Topuria, C. Nyffeler, K. Virwani, A. Padilla, R.M. Shelby, M. Eskandari, B. Jackson, B.-S. Lee, Observation and modeling of polycrystalline grain formation in $Ge_2Sb_2Te_5$. J. Appl. Phys. **111**, 104308 (2012)
52. N. Ciocchini, E. Palumbo, M. Borghi, P. Zuliani, R. Annunziata, D. Ielmini, Modeling resistance instabilities of set and reset states in phase change memory with Ge-rich GeSbTe. IEEE Trans. Electron Devices **61**, 2136–2144 (2014)
53. T. Siegrist, P. Jost, H. Volker, M. Woda, P. Merkelbach, C. Schlockermann, M. Wuttig, Disorder-induced localization in crystalline phase-change materials. Nat. Mater. **10**, 202–208 (2011)
54. D. Ielmini, D. Sharma, S. Lavizzari, A.L. Lacaita, Reliability impact of chalcogenide-structure relaxation in phase change memory (PCM) cells – Part I: Experimental study. IEEE Trans. Electron Devices **56**, 1070–1077 (2009)
55. N. Ciocchini, M. Cassinerio, D. Fugazza, D. Ielmini, Modeling of threshold voltage drift in phase change memory (PCM) devices. IEEE Trans. Electron Devices **59**, 3084–3090 (2012)
56. D. Ielmini, Y. Zhang, Evidence for trap-limited transport in the subthreshold conduction regime of chalcogenide glasses. Appl. Phys. Lett. **90**, 192102 (2007)
57. D. Fugazza, D. Ielmini, S. Lavizzari, A.L. Lacaita, Distributed-Poole-Frenkel modeling of anomalous resistance scaling and fluctuations in phase-change memory (PCM) devices. IEDM Tech. Dig., 723–726 (2009)
58. S.C. Agarwal, Role of potential fluctuations in phase-change GST memory devices. Phys. Status Solidi B **249**, 1956–1961 (2012)
59. A.K. Jonscher, J. Phys. C **4**, 1331 (1971)
60. A.K. Jonscher, C.K. Loh, J. Phys. C **4**, 1341 (1971)
61. A.K. Jonscher, R.M. Hill, Electrical conduction in disordered nonmetallic films, in *Physics of Thin Films*, G. Hass, M. H. Francombe, R. W. Hoffman (Eds), (Academic, New York, 1975), p. 8
62. S. Tam, P.-K. Ko, C. Hu, Lucky-electron model of channel hot-electron injection in MOSFET'S. IEEE Trans. Electron Devices **31**, 1116–1125 (1984)

63. S. Lavizzari, D. Ielmini, A.L. Lacaita, Transient simulation of delay and switching effects in phase change memories. IEEE Trans. Electron Devices **57**, 3257–3264 (2010)
64. D. Ielmini, D. Mantegazza, A.L. Lacaita, A. Pirovano, F. Pellizzer, Parasitic reset in the programming transient of phase change memories. IEEE Electron Device Lett. **26**, 799–801 (2005)
65. S. Lavizzari, D. Sharma, D. Ielmini, Threshold-switching delay controlled by 1/f current fluctuations in phase-change memory devices. IEEE Trans. Electron Devices **57**, 1047–1054 (2010)
66. S. Lavizzari, D. Ielmini, D. Sharma, A.L. Lacaita, Transient effects of delay, switching and recovery in phase change memory (PCM) devices. IEDM Tech. Dig, 215–218 (2008)
67. M. Cassinerio, N. Ciocchini, D. Ielmini, Logic computation in phase change materials by threshold and memory switching. Adv. Mater. **25**, 5975–5980 (2013)
68. D. Loke, J.M. Skelton, W.-J. Wang, T.-H. Lee, R. Zhao, T.-C. Chong, S.R. Elliott, Ultrafast phase-change logic device driven by melting processes. Proc. Natl. Acad. Sci. U. S. A. **111**, 13272–13277 (2014)
69. Y. Li, Y.P. Zhong, Y.F. Deng, Y.X. Zhou, L. Xu, X.S. Miao, Nonvolatile “AND,” “OR,” and “NOT” Boolean logic gates based on phase-change memory. J. Appl. Phys. **114**, 234503 (2013)
70. D. Kuzum, R.G.D. Jeyasingh, B. Lee, H.-S.P. Wong, Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. Nano Lett. **12**, 2179 (2012)
71. G.W. Burr, R.M. Shelby, S. Sidler, C. di Nolfo, J. Jang, I. Boybat, R.S. Shenoy, P. Narayanan, K. Virwani, E.U. Giacometti, B.N. Kurdi, H. Hwang, Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element. IEEE Trans. Electron Devices **62**, 3498 (2015)
72. S. Ambrogio, N. Ciocchini, M. Laudato, V. Milo, A. Pirovano, P. Fantini, D. Ielmini, Unsupervised learning by spike timing dependent plasticity in phase change memory (PCM) synapses. Front. Neurosci. **10**, 56 (2016)
73. D. Ielmini, S. Lavizzari, D. Sharma, A.L. Lacaita, Physical interpretation, modeling and impact on phase change memory (PCM) reliability of resistance drift due to chalcogenide structural relaxation. IEDM Tech. Dig., 939–942 (2007)
74. A. Pirovano, A.L. Lacaita, F. Pellizzer, S.A. Kostylev, A. Benvenuti, R. Bez, Low-field amorphous state resistance and threshold voltage drift in chalcogenide materials. IEEE Trans. Electron Devices **51**, 714 (2004)
75. D. Ielmini, A.L. Lacaita, D. Mantegazza, Recovery and drift dynamics of resistance and threshold voltages in phase change memories. IEEE Trans. Electron Devices **54**, 308 (2007)
76. S. Gabardi, S. Caravati, G.C. Sosso, J. Behler, M. Bernasconi, Microscopic origin of resistance drift in the amorphous state of the phase-change compound GeTe. Phys. Rev. B **92**, 054201 (2015)
77. S. Roorda, W.C. Sinke, J.M. Poate, D.C. Jacobson, S. Dierker, B.S. Dennis, D.J. Eaglesham, F. Spaepen, P. Fuoss, Structural relaxation and defect annihilation in pure amorphous silicon. Phys. Rev. B Condens. Matter **44**(8), 3702–3725 (1991)
78. K.P. Chik, S.Y. Feng, S.K. Poon, Generation of dangling bonds by high temperature annealing and hopping conduction in amorphous silicon films. Solid State Commun. **33**(10), 1019–1023 (1980)
79. E.P. Donovan, F. Spaepen, D. Turnbull, J.M. Poate, D.C. Jacobson, Calorimetric studies of crystallization and relaxation of amorphous Si and Ge prepared by ion implantation. J. Appl. Phys. **57**(6), 1795–1804 (1985)
80. L.L. Snead, S.J. Zinkle, Structural relaxation in amorphous silicon carbide. Nucl. Instrum. Methods Phys. Res. B Beam Interact. Mater. At. **191**(1–4), 497–503 (2002)
81. N.M.J. Conway, A. Ilie, J. Robertson, W.I. Milne, A. Tagliaferro, Reduction in defect density by annealing in hydrogenated tetrahedral amorphous carbon. Appl. Phys. Lett. **73**(17), 2456–2458 (1998)

82. M. Boniardi, D. Ielmini, Physical origin of the resistance drift exponent in amorphous phase change materials. Appl. Phys. Lett. **98**, 243506 (2011)
83. M. Boniardi, D. Ielmini, S. Lavizzari, A.L. Lacaita, A. Redaelli, A. Pirovano, Statistics of resistance drift due to structural relaxation in phase-change memory arrays. IEEE Trans. Electron Devices **57**, 2690–2696 (2010)
84. A. Sebastian, N. Papandreou, A. Pantazi, H. Pozidis, E. Eleftheriou, Non–resistance–based cell–state metric for phase–change memory. J. Appl. Phys. **110**(08), 084505 (2011)
85. S. Kim, N. Sosa, M. BrightSky, D. Mori, W. Kim, Y. Zhu, K. Suu, C. Lam, A phase change memory cell with metallic surfactant layer as a resistance drift stabilizer. IEDM Tech. Dig, 762 (2013)

Chapter 3
Thermal Model and Remarkable Temperature Effects on the Chalcogenide Alloy

Mattia Boniardi

3.1 Phase Change Memory Electrothermal Model Fundamentals

As already introduced in the previous chapters, the phase change memory (PCM) is based on the unique properties of the chalcogenide material, usually a GeSbTe (GST) compound, able to make reversible phase transitions between an amorphous and a polycrystalline state, owing to the combined effect of both the electronic switching and the memory switching effects [1]. A detailed discussion on this point will be reported in Chap. 4. In both operations, the crystal-to-amorphous transition (SET to RESET) and vice versa (RESET to SET), the starting point toward the phase change mechanism is the capability of raising temperature within the chalcogenide layer as the result of a program current flowing in a heater electrode in series and in contact to GST and/or thanks to a current crowding effect into a properly shaped cell able to allow the self-heating of GST.

Among the most relevant parameters to describe the electrothermal efficiency of PCM devices are the melting current, I_{MELT}, and the SET resistance, R_{SET}, of the PCM cell, the former representing power consumption and the latter representing the material conductance in the SET state. Since in the real device the SET state is the detected one, namely, the state in which the PCM cell conducts current, such state intrinsically determines the device read latency: the lower the SET state resistance, the higher the current sensed by the sense amplifier and, eventually, the faster the sensing (readout) operation or read latency. Consumption and latency are intrinsically linked owing to the *Joule effect* [2]: in general, the higher the R_{SET}, the lower the I_{MELT} required to allow GST reach the molten state and the higher the

M. Boniardi (✉)
Micron Semiconductor Italia s.r.l., Vimercate 20871, Italy
e-mail: mboniardi@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_3

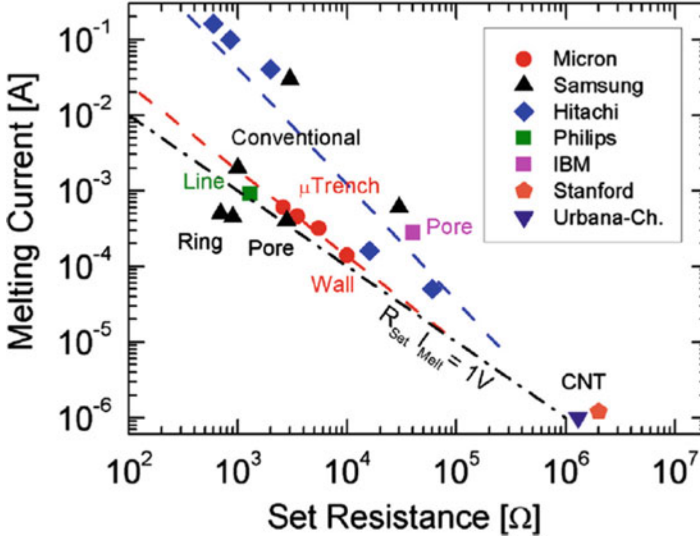


Fig. 3.1 Melting current as a function of SET resistance in different PCM cell architectures from different references [3]. Data include different cell architectures and different geometries

read latency. Similarly, the lower the R_{SET}, the higher the I_{MELT}, leading to critical power consumption and program throughput but decreased read latency.

This anticorrelation does have some hidden parameters worth of a comment: in fact, a change in R_{SET} can be associated to a geometry change, e.g., a variation of the cell critical dimension (CD), likely leading to some possible melting current and power consumption modulation induced by a modulation of the cell overall thermal resistance, R_{TH}, and of the cell ON-state resistance, R_{ON}. Figure 3.1 reports the I_{MELT}–R_{SET} correlation plot for some cell architectures [3] that can be found in the literature. The best working area for PCM would be the left-bottom hand side of the plot, that is, both low R_{SET} and I_{MELT}, or minimized $R_{SET}{\cdot}I_{MELT}$ product. The most relevant equations describing the PCM working principle are a condensed-parameter, simplified *Fourier equation* for heat transfer and the *Joule heating* equation:

$$\Delta T = R_{TH} \cdot P \tag{3.1}$$

$$P = R_{ON} \cdot I^2 \tag{3.2}$$

In Eq. (3.1) $\Delta T = T – T_{ROOM}$, R_{TH} is the cell overall thermal resistance and P is the power injection required to obtain ΔT, according to the simplified *Fourier equation*. In Eq. (3.2) R_{ON} is the ON-state resistance of the cell, in the *Joule heating* context. It is worth considering that R_{ON} includes contributions from GST and possible electrodes in series [4] and that the resistivity of GST in the over-melting condition lies in the range of $3 \cdot 10^{-4}$–$5 \cdot 10^{-4}$ Ωcm, owing to a resistivity drop between the under-melting and over-melting conditions [5]. By combining

Eqs. (3.1) and (3.2), the found ingredients to have a temperature rise within the phase change material, for a given I, are the thermal resistance and the ON-state resistance of the cell, depending on both GST and electrode ones and the injected current, like the I_{MELT} as a minimum boundary. Hence, $\Delta T = f(R_{TH}, R_{ON}, I)$. Both R_{TH} and R_{ON} are architecture, technology node, and material dependent and both correlate with R_{SET}. This provides stronger basis to the plot in Fig. 3.1. Thus, among the most important parameters for PCM efficiency, we list here R_{ON} and R_{TH}, both responsible of a temperature rise occurring in the material. The former is very easy to evaluate, that is, by fitting the pulsed I-V curve of the PCM cell at high field (ON-state); the latter is nontrivial to be extracted. The next paragraph will be devoted to its evaluation on the real device.

3.2 Fourier's Law and Thermal Resistance Experimental Estimation

The capability of experimentally estimating the temperature, in the PCM cell during the phase change operation, represents a milestone related to the comprehension of the phase change mechanism: in fact, it may provide insights into the thermal efficiency of different cell architectures, in general, and it may enable a new focus on geometry and on the phase change material. The cell thermal parameters could be calculated through electrothermal simulations [6] or, in this case, experimentally evaluated thanks to custom electrical measurements. As already mentioned, the phase change operation relies on the capability to effectively translate an electrical power injection in the PCM cell, into a coherent temperature rise, owing to the *Fourier's law* of heat transfer; hence, an experimentally driven modulation of power injection can reveal the PCM cell thermal response. This paragraph will be devoted to the presentation of such experimental techniques.

3.2.1 Thermal Resistance and Temperature Estimation

The analysis of the resistance–current (R–I) characteristic curves of the PCM cell can highlight a path to the development of an experimental method for cell internal temperature evaluation, which could be applied to make scaling and geometry and composition considerations for the PCM technology. Further, the capability of estimating the cell internal temperature may be a way of characterizing the properties of different phase change materials, as integrated in PCM cells with the same geometry [7] and, likely, of optimizing the phase change material composition. The developed method is based on the study of the resistance–power ($R-P$) curves taken on PCM cells, as shown in Fig. 3.2. The quantification of power injection, P, into the cell, and the respective increase in resistance after a delivered pulse can be

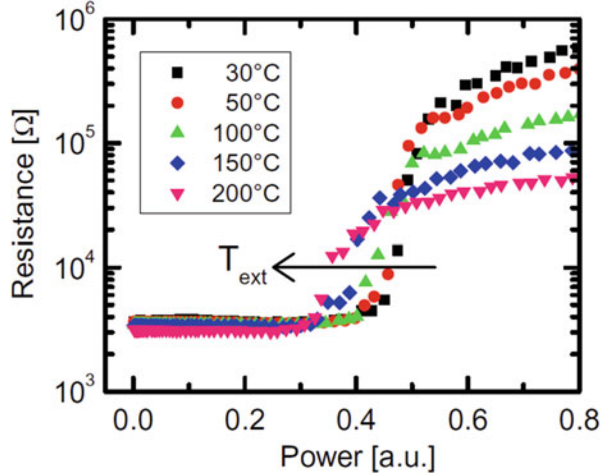


Fig. 3.2 Resistance–power curves as a function of environment temperature. The SET to RESET edges are shown to shift to lower values of power with increasing temperature. The sensitivity is represented by the cell thermal resistance

obtained by combining the measurements of the programming curves ($R-I$) and pulsed characteristic curves ($I-V$), according to the relation $P = V \cdot I$. Moreover, the $R-P$ data reported in Fig. 3.2 are obtained at different environment temperatures, namely, $T = 30$ °C, 50 °C, 100 °C, 150 °C, and 200 °C, forced by a thermal chuck on which the wafer with sample cells is fixed. The highest temperature needs to be carefully evaluated in order to avoid the amorphous crystallization at the RESET side of the $R-P$ plot: in this case, $T = 200$ °C is considered as a safe condition, and no resistance drop was registered in the time scale from the end of the RESET pulse and the resistance DC measurement (1 s range). The different environment temperature results into three different effects on the experimental $R-P$ curves.

First, an increasing temperature results in a decreasing RESET resistance, according to the conduction activation, namely, $E_C = 0.3$ eV [7, 8] for the full RESET state (full amorphous phase of GST); second, an increasing temperature results in an only slightly decreasing SET resistance, due to the slighter conduction activation of crystalline GST, namely, $E_C \approx 0.02$ eV [8]. Such behavior has been discussed in Chap. 2 also. Third, and most important, Fig. 3.2 highlights a shift of the SET to RESET edge of the programming curves to lower power for increasing the environment temperature, T. Indeed, the higher the T, the lower the power injection required for the electrically driven melting of GST, as a consequence of a temperature rise enabled by the environment (thermal chuck) contribution and by power injection. So, by expanding the ΔT term of Eq. (3.1), the following equation can be written and considered as representative of the experimental path to be followed for internal temperature and thermal resistance extraction at the melting condition:

$$T_{\mathrm{MELT}} = T_{\mathrm{ROOM}} + R_{\mathrm{TH}} \cdot P_{\mathrm{MELT}}(T_{\mathrm{ROOM}}) \tag{3.3}$$

where T_{ROOM} is the temperature boundary condition forced by the thermal chuck, T_{MELT} is the cell internal temperature during a specific programming operation (here the melting condition), and R_{TH} is the thermal resistance of the cell, with the additional hypothesis that it is not dependent on T_{ROOM}. Equation (3.3) is simply Eq. (3.1) applied to this particular problem and to the particular onset of phase change, i.e., the melting condition. The cell internal temperature for the onset of phase change is $T_{MELT} \leq T \leq T_{RESET}$ where T_{RESET} is referred to as the temperature within the cell core when the RESET condition is reached. Consider now a particular device status, e.g., the material melting condition, defined as the buildup of an amorphous thin film on the heater–GST interface and production of the very first resistance increase from the SET state. In this case, Eq. (3.3) holds true and T_{MELT} is thus reached owing to the superposition of a forced T_{ROOM} and an electrical power injection contribution; equivalently, a forced T_{ROOM} variation results into the modulation of the required electrical power to reach the material melting point with a fixed thermal resistance. The dissipated power at melting can be thus evaluated from the SET to RESET edge for each curve. So, in principle, n equations could be written as follows:

$$T_{MELT} - T_{ROOM,1} = R_{TH} \cdot P_{MELT,1}(T_{ROOM,1}) \quad (3.4a)$$
$$T_{MELT} - T_{ROOM,2} = R_{TH} \cdot P_{MELT,2}(T_{ROOM,2}) \quad (3.4b)$$
$$T_{MELT} - T_{ROOM,n} = R_{TH} \cdot P_{MELT,n}(T_{ROOM,n}) \quad (3.4c)$$

in which T_{MELT} is the cell internal temperature in the considered state. Subtracting the left- and right-hand sides, respectively, of any couple of Eqs. (3.4), it is easy to rule $\Delta T_{ROOM} = -R_{TH} \cdot \Delta P_{MELT}$. The extracted power values can be arranged within an external temperature–power ($T-P$) diagram, as shown in Fig. 3.3. The y-axis reports the externally forced temperature, T_{ROOM}, and the x-axis accounts for the power extracted from curves in Fig. 3.2 at variable T_{ROOM}, in the considered melting condition and in the hypothesis that R_{TH} is not a function of T_{ROOM}. Such hypothesis is supported by data in Fig. 3.3, pointing out a good linearity of the

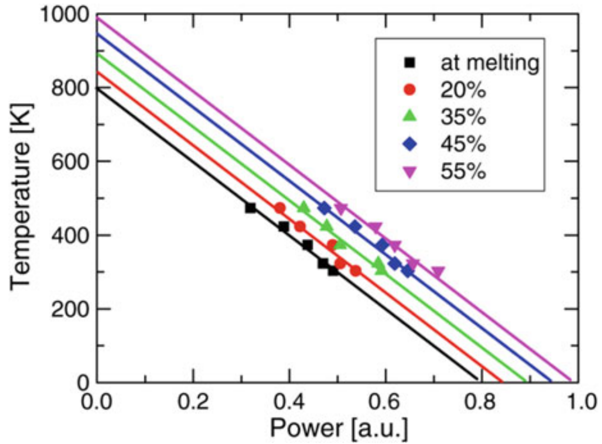


Fig. 3.3 Room temperature as a function of power injection to obtain fixed amorphous volumes from the plots in Fig. 3.2

extracted points (at melting) in the observed temperature range. The described power extraction could be applied exactly in the same way as developed in the melting condition, to other states in the SET to RESET edge. To this purpose, to rule out the impact of the amorphous activation with temperature on the resistance readout, a resistance state, e.g., the RESET state resistance, should be taken as reference, representing a single phase change volume of GST among the experiments made at different T_{ROOM}. Following this approach, each plot in Fig. 3.2 could be normalized to this value, as better explained in Ref. [9]. On one hand, the data points in Fig. 3.3 can be modeled by a linear behavior, supporting the validity of Eq. (3.3). The slope of the fitting curve represents the effective thermal resistance of the PCM cell, accounting for all the possible directions for thermal flux propagation toward the thermal boundaries and considering the possibility of a nonideal power generation peak at the heater–GST interface. On the other hand, the intercepts of Eq. (3.3), at $P = 0$ (zero electrical power injection), represent an approximated extraction of the cell internal temperature during the phase change operation, required to obtain a first rise in resistance of SET to RESET curves in Fig. 3.2. The zero-power T intercepts are based on the assumption that, at zero electrical power injected into the PCM cell, the whole amount of temperature required for the onset of phase change should be entirely provided externally by the experimental setup (thermo-chuck), namely, $T_{ROOM} = T_{MELT}$, from Eq. (3.3). Hence, this method, introduced for the first time by Ref. [9] and also adopted in Ref. [10], allows for an indirect internal temperature measurement during the PCM operation, with extracted values of temperature very close to those expected from the GST melting temperature on, namely, $T_{MELT} \geq 880$ K [7]. A lower melting temperature extracted by the black points could be ascribed to an actually lower melting temperature occurring in the very first amorphous volume established owing to a Sb-rich alloy in this region, as will be reported in the following paragraph.

It is worth noticing that the R states below 55% of the SET to RESET transient should be considered only, in order to avoid incorrect values of resistance due to resistance drift at high R that can vary with temperature, thus distorting the relationship between R and the amorphous thickness. Electrothermal simulations of the 90 nm cell have been also reported [9], showing a good agreement between the simulated temperature–voltage ($T-V$) diagram and the data points extracted from the $T-P$ curves in Fig. 3.3.

The simulated $T-V$ plots [9] allow to predict the cell internal temperature at the full reset condition, namely, $T = 1150$–1200 K, that is the highest temperature reached at the heater–GST interface during the phase change operation. This represents useful information at the device level in order to support the modeling of reliability effects such as the possible endurance related issues and proximity (thermal) disturb.

3.2.2 *Impact of Geometry on the Cell Thermal Resistance*

The PCM cell geometry shows some impact in the cell thermal resistance and represents one of the possible knobs to get better heat confinement. A scaling analysis between two cells, a 90 nm node cell and a shrunk version of this one, is reported in Fig. 3.4a: the different slope of the $T-P$ diagrams taken for the two cells represents the difference between their thermal resistances and correctly accounts for the aggressive scaling factor applied ($k \approx 4$). In this case, the thermal resistance R_{TH} scaling is consistent with the shrink of the critical dimension by about the same factor, as extracted from Fig. 3.4a, $k \approx 4$. Moreover, data fitting curves show about the same intercept on the temperature axis, that is, close to the melting temperature of GST, which is used as phase change layer in both cells with different geometry, with the same composition ($Ge_2Sb_2Te_5$). This further enforces the reliability of the thermal resistance extraction method, and the melting temperature estimation, as extracted from the intercept of the fitting curves in Fig. 3.4a. Electrical power shows a decreasing fashion between the two cells due as well to the smaller contact area between the heater electrode and the GST layer introduced by scaling the critical dimension and hence increasing the ON-state resistance, R_{ON}. By adopting the same method and by considering the comments done on Fig. 3.1, thermal resistance can be reported for cells belonging to different PCM technology nodes, namely, 180 nm [11], 90 nm [12], and 45 nm [13], with their relative variations of I_{MELT} and R_{SET}. Figure 3.4b points out a decreasing I_{MELT} with increasing R_{SET}, along the different technology nodes; here a different technology node means different geometry in general: the geometry change is not essentially a critical dimension change but also a difference in other features (thickness, alloy). By the way, a decreasing I_{MELT}, increasing R_{SET}, is shown to be in trend with an increasing overall R_{TH}. Thanks to the same procedure, the thermal resistance can be extracted as a function of GST thickness, L_{GST}, and heater length, L_H, here reported on the 45 nm technology node, enabling the definition of some guidelines for architecture optimization in terms of geometry [2]. Figure 3.5a takes the comparison of cells

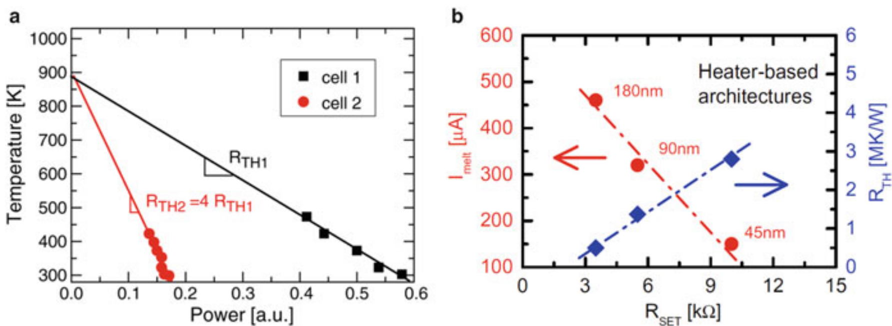


Fig. 3.4 Temperature–power plot of a 90 nm cell and its shrunk version with scaling factor $k = 4$ (**a**) and $I_{MELT}-R_{SET}$ plot for three PCM technology nodes, with thermal resistance extraction (**b**)

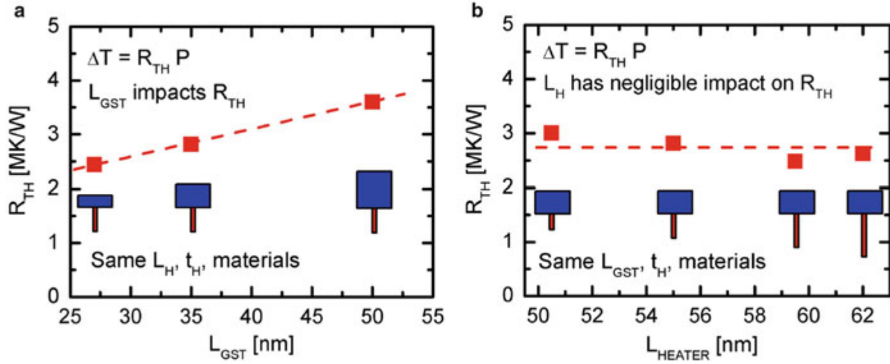


Fig. 3.5 Thermal resistance extraction for different GST length or thickness (**a**) and for different heater length (**b**) on the *wall* PCM cell at the 45 nm technology node

with different L_{GST}, namely, 50 nm, 35 nm, and 27 nm, in terms of R_{TH}: a decreasing L_{GST} results into a decreasing cell thermal resistance, R_{TH}, in agreement with a reduced heat path toward the top electrode contact (TEC) for thinner L_{GST}. Indeed, the thinner the chalcogenide film, the closer the top thermal boundary to the cell active volume.

Moreover, such experimental evidence is in agreement with thermal disturb evidences: an increasing cycle capability on an aggressor cell before data retention failure on the closest victim has been reported with thinner GST, with respect to a thicker GST reference, further supporting the model of a shorter heat escape path toward the TEC, preventing/mitigating a lateral-bound temperature rise [14]. On the other hand, trials realizing different heater lengths, L_{H}, do not introduce significant shift of the R_{TH} extraction: in fact a heater length modulation in the range 50–62 nm is reported to provide aligned thermal resistance values, with R_{TH} slightly varying in the range 2.5–3 K/μW, as reported in Fig. 3.5b, for the *wall* 45 nm technology node. This suggests that the preferential path for heat propagation in the *wall* architecture involves mainly the distance between the heater–GST interface and the top thermal boundary represented by the TEC, namely, the path through the thickness of the GST film toward the top thermal boundary. The heat path from the heater–GST interface downward to the BEC is, instead, not preferential, likely due to its higher thermal resistance, associated with the thickness and electrical/thermal resistivity of the heater electrode. Hence, the parallel heat path from the cell core toward the TEC and the BEC results into a simplified model in which the overall thermal resistance of the cell is dominated by the heat path toward the TEC.

3.2.3 Impact of GeSbTe Alloy on the Cell Thermal Resistance

The GeSbTe composition of the chalcogenide alloy can be also responsible of different thermal resistance of the PCM cell. Such idea has been suggested by some trials aiming at reducing power injection at melt thanks to the use of a GST composition featuring a lower melting temperature. In fact, by considering the GeSbTe ternary diagram, the composition could be moved away from the most used $Ge_2Sb_2Te_5$, toward a Te-poor, Sb-/Ge-rich region, featuring, according to data in the literature [15, 16], a lower melting temperature, T_{MELT}. The investigation of possible thermal resistance mismatch depending on the specific composition was boosted by the experimental evidence that Te-poor alloys are characterized by higher current to obtain the melting condition of GST, despite the reported lower T_{MELT} with respect of Te-rich GeSbTe alloys. The prime suspects of such behavior are the ON-state resistance and the thermal resistance of the cells with different alloys. The former can be discarded since cell geometries are the same, their SET resistance and high field resistance do not change; in order to investigate the latter, the same method explained in the first paragraph was adopted. Figure 3.6a reports indeed the melting current/power behavior upon composition, highlighting an increase of I_{MELT}/P_{MELT} toward the Te-poor direction, despite a decreasing melting temperature, from 870 K to about 800 K [15, 16], as highlighted by the segment in the GeSbTe ternary diagram with T_{MELT} as contour color (Fig. 3.6b). Thus, the analysis of the resistance–power ($R-P$) characteristic curves of the studied alloys, taken at different environment temperature T_{ROOM}, was done on the same cell geometry, revealing that different alloys can impact the cell thermal resistance. The $T-P$ diagram in Fig. 3.7 allows for the extraction of both the thermal resistance R_{TH}, as the melting-power P_{MELT} sensitivity on T_{ROOM}, and the melting temperature T_{MELT} at zero power injected, as in the previous cases. Results are reported in Fig. 3.7 for two alloys with different Te atomic percentage (at. %) concentrations, namely, 50% and 35%. The plot evidences on one hand a lower melting

Fig. 3.6 Current/power variation upon Te atomic percentage of GST (**a**) and melting temperature contour plot of the GeSbTe ternary diagram (**b**)

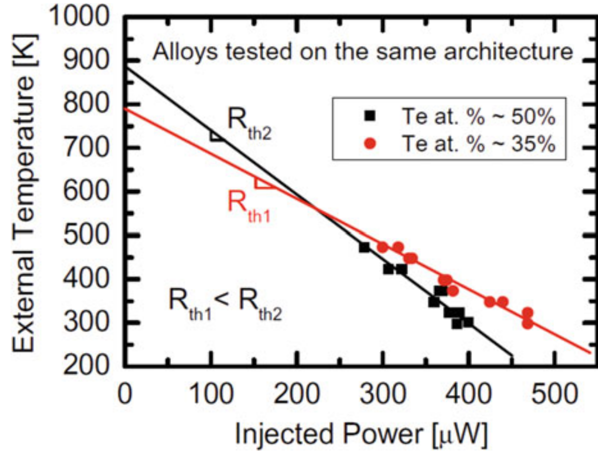


Fig. 3.7 Room temperature–power plot for a Te-rich and a Te-poor alloy, on the same cell geometry, to check the thermal resistance behavior of different alloys

temperature extracted for the Te-poor compound, around 800 K, as expected from Fig. 3.6b and on the other hand a lower R_{TH} in the Te-poor alloy ($R_{TH1} < R_{TH2}$). This confirms that despite Te-poor GST having a lower melting temperature, the material itself is responsible of a worse thermal confinement during the phase change operation thus requiring a slightly higher injection of power to reach its melting condition than its Te-rich counterpart, showing instead a higher R_{TH} that is a better thermal confinement.

3.3 Temperature-Related Effects on the Phase Change Material

As already mentioned along this chapter and on the previous ones, the phase change material is submitted to high magnitude stress in terms of electrical and thermal physical quantities. In fact high electrical fields and strong thermal gradients are the main ingredients for the phase change operation. It is the direct aim of this chapter to get the reader aware of the fact that temperature plays a capital role in determining the properties of the phase change material (GST or other) and that temperature, along with other stresses (current density, electric field), pushes the phase change material into a nonuniform working condition. Such working conditions should be taken into account in order to promote a correct modeling of the phase change material through the device cycle lifetime and a forward-looking treatment of the endurance-/fatigue-related failures.

3.3.1 Fragility Behavior of Chalcogenide Alloys for PCM

Crystallization represents a very important chapter for phase change memory and is one of the most critical mechanisms, since it is involved in the SET operation performance and hence is a major contributor of the program latency. Crystallization deals with the reorganization of atoms, starting from a liquid or an amorphous phase into a long-range ordered structure. Thanks to the absence of local composition change during crystallization and to the presence of rock salt structures already in the liquid that can undergo reorientation owing to the high concentration of vacancies, crystallization is reported to be a fast phenomenon. As a counterpart, GST, or any other phase change alloy suitable for PCM, should feature high data retention capability: the amorphous phase should be sufficiently stable to retain the RESET state under the environment temperature. During the SET operation, crystallization takes place in the supercooled liquid regime, above T_g, namely, the glass transition temperature of GST, and below T_{MELT}. Thus, a parameter of capital importance is the sensitivity of liquid structure to temperature changes, namely, the temperature dependence of the atomic mobility in the supercooled liquid or viscosity, η.

In fact, liquids can be classified according to the viscosity behavior as a function of temperature, after the definition by Angell [17, 18]: on one hand, strong liquids show a temperature behavior of viscosity quite close to the Arrhenius one in the whole temperature range over T_g; on the other hand, in fragile liquids, viscosity shows a marked non-Arrhenius behavior with temperature that could be summarized by an effective activation energy of viscosity, being higher close to T_g and strongly decreasing when temperature is raised toward T_{MELT}. The fragility of a glass-forming liquid can be defined as:

$$m = \left. \frac{d(Log\eta)}{d(T_g/T)} \right|_{T=T_g} \tag{3.5}$$

The evaluation of m is done at $T = T_g$, where $Log_{10}\,\eta = 10^{12}$ Pa s, by definition. A viscosity plot in the supercooled range as a function of temperature, scaled by the particular values of T_g, is widely known in the literature as the *Angell plot* and has been reported for many glass-forming liquids, covering the spectrum from strong to fragile liquids: SiO_2 (strongest liquid represented), As_2S_3, $Na_2O.2SiO_2$, glycerol, $Te_{85}Ge_{15}$, and o-terphenyl (most fragile liquid–organic material) [17, 18]. An *Angell* plot is reported in Fig. 3.8. For example, open network liquids, like SiO_2 or GeO_2, show a marked Arrhenius variation of viscosity/structural relaxation time between T_g and the high temperature limit, leading to the strong liquid pattern of the plot; on the other hand, other liquids characterized by simple nondirectional coulomb attractions lead to the fragile behavior. Fragile liquids have glassy state structures able to be reorganized over a wide variety of different particle

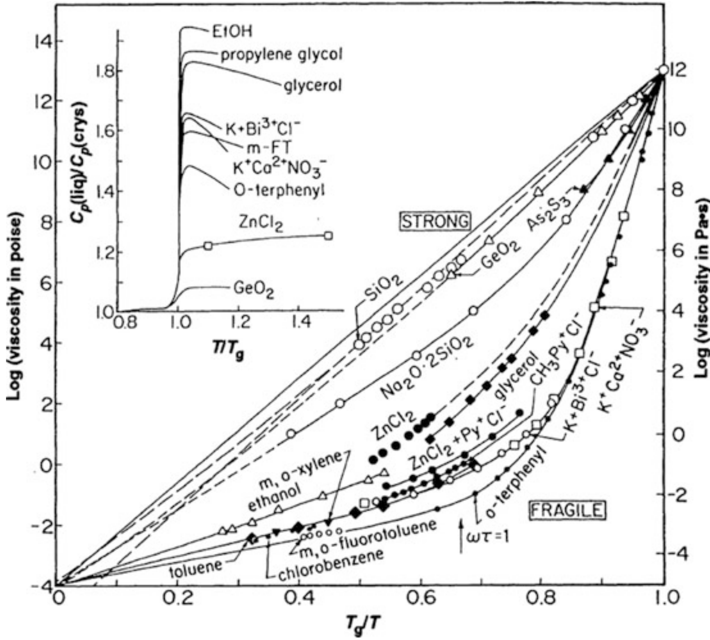


Fig. 3.8 *Angell* plot of Log (viscosity) as a function of normalized temperature

orientations and coordination states. Strong liquids, instead, show a built-in resistance to structural change and show little reorganization over a wide variation of temperature. The fragility of the most studied chalcogenide material for phase change memory applications, $Ge_2Sb_2Te_5$, has been measured and reported, for the first time, by J. Orava [19, 20] thanks to ultrafast differential scanning calorimetry (DSC) characterizations. The main difference of the ultrafast method, with respect to conventional calorimetry, is the capability of heating up the sample with very high heating rates, Φ, up to $4 \cdot 10^4$ K s^{-1}, allowing for the exploration of crystallization exothermal peaks at much higher values of temperature with respect to conventional calorimetry. In fact the temperature range results enlarged up to $T_p = 650$ K, bridging the gap 520–650 K, unexplored by previous literature. The product of such characterization could be represented thanks to the *Kissinger* plot [18], that is, a plot of $\ln(\Phi/T_p^2)$ versus $1/T_p$, as reported in Fig. 3.9. This plot is expected to provide a gradient of $-Q/R$, where Q is the activation energy of crystallization and R is the gas constant.

Orava et al. comment that since in conventional DSC the heating rate range is very narrow, straight plots of $\ln(\Phi/T_p^2)$ versus $1/T_p$ can be obtained, thus leading to a single activation energy of crystallization. Instead, data obtained with ultrafast

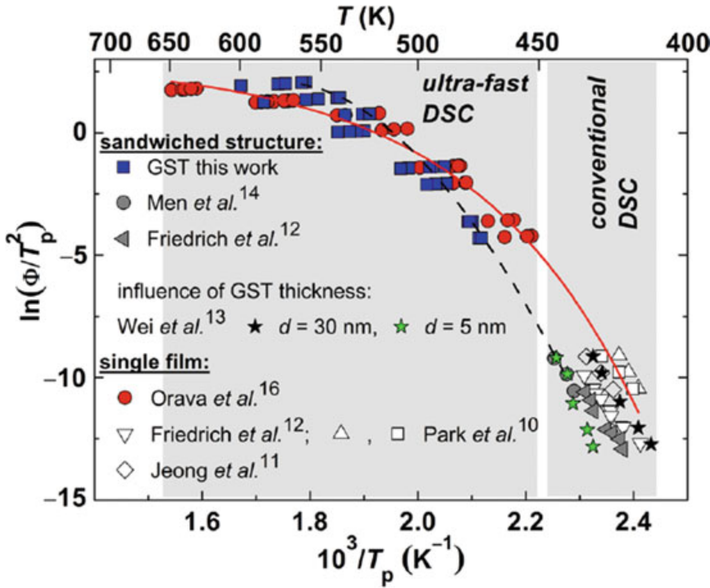


Fig. 3.9 Kissinger plot with GST points: the ultrafast calorimetry technique introduced by Orava and co-workers allowed the extension of the observable range

DSC can be considered as a good extension of the observable range of conventional DSC, allowing for a curved plot, that is, an effective activation energy decreasing at higher temperature, typical of fragile liquids. From the *Kissinger* plot, the crystallization velocity can be carried out as the product of a kinetic and a thermodynamic contribution [19]. Reasonably assuming that the kinetic contribution is inversely proportional to viscosity, η, close to the melting temperature, T_{MELT}, the *Kissinger* plot can be translated into the *Angell* plot describing viscosity [19]. However, by pushing temperature below T_{MELT}, toward T_g, the viscosity plot shows a behavior only roughly matching the measured viscosity data, pointing out that, for fragile liquids, the inverse proportionality assumption between η and the kinetic contribution cannot fit the expected behavior, leading to the breakdown of the Stokes–Einstein relation. Orava and co-workers point out that the kinetic coefficient u_{kin} for the crystallization rate of GST is thus progressively decoupled from viscosity, η, as temperature is lowered toward T_g, with the following scaling law: $u_{kin} = B \cdot \eta^{-\xi}$ with $\xi \approx 0.67$. The gradient of this fit line as $T_g/T \rightarrow 1$ gives a fragility value $m = 90$ [19, 20].

Thus, in terms of fragility, liquid GST is very close to several organic liquids and shows a strong temperature dependence of the crystallization rate, leading to very fast crystallization at high temperature. In doing such study, Orava et al. were guided by the seminal work by Ediger and co-workers, establishing, for a wide range of liquids, a proportionality between u_{kin} and $\eta^{-\xi}$, with the exponent $\xi < 1$ and, most interestingly, strongly correlated with the fragility of the liquid [21], as represented in Fig. 3.10 for organic liquids, metallic glasses, and oxides. In

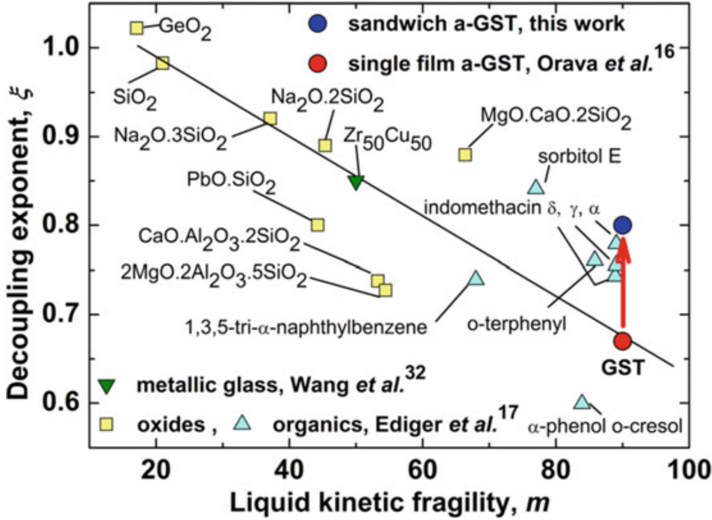


Fig. 3.10 Correlation plot between the growth velocity/viscosity decoupling exponent and fragility coefficient. The more fragile a liquid, the higher the decoupling exponent of the power law

particular, the ξ values are approximately represented by a linear correlation with fragility, namely, $\xi \approx 1.1–0.005\ m$. So, since fragility is explicitly a property of the supercooled liquid only, the authors claim that liquid has no memory/knowledge of the crystal state, suggesting that u_{kin} is largely determined by the dynamics of the liquid alone [21]. As fragile liquids are cooled down toward T_g, local relaxation is expected to occur at different rates and at different places within the liquid, that is, the dynamics is spatially heterogeneous [21]. Such dynamical heterogeneities have been object of recent theoretical investigations [22]. Some authors depict this phenomenology as clearly distinguishable regions of the studied volume, with clearly different relaxation rates; actually no experiment could determine whether they are distinguishable islands or they could be represented as a continuous evolution of relaxation rates from region to region. Then the decoupling and marked non-Arrhenius behavior of u_{kin} and η show that conventional crystallization measurements close to T_g have limited relevance. The fragility context represents a very powerful instrument for the comprehension of the very fast crystallization behavior of phase change alloys.

3.3.2 Segregation Behavior of Chalcogenide Alloys for PCM

As already highlighted, PCM are enabled by a temperature rise in the active core of an electro-thermal device. The onset of phase change requires temperature to be raised to or over the melting temperature of GST. Beyond temperature, and the onset of possible temperature gradients within GST, PCM operation is enabled by a

current flowing through the device, on its turn, enabled by an electrical field applied to the cell. All these solicitations do result into cell operation with a nonuniform alloy. Recently researchers have pushed a lot of focus toward the investigation of the phase change composition as a function of the number of programming cycles at different conditions and by using different cell architectures. In fact the comprehension of the mechanisms involved in endurance is crucial since those are linked with the material stability [23] after electrical/thermal stress during the device lifetime. Hence, the focus on elemental migration in PCM as a function of an applied stress is being pursued for researching the root causes of both the electrical parameters alteration during cycling, like R_{SET}, I_{MELT}, and R_{RESET}, [24] and the PCM failure mechanisms.

Rajendran et al. reported that a material change, suggested by the deviation of the RESET current and the programmed resistance levels on cycling, is confirmed by EDX analyses pointing out a Sb-richer and Ge-depleted volume in the cell core [23]. No clear variations on Te concentration could be reported between the fresh and cycled cases. A Sb-rich region close to the bottom electrode contact (BEC) is claimed to increase the contact area between BEC and GST, like having a BEC protruded into GST), explaining the steady decrease in SET resistance with cycling, as reported in Fig. 3.11. Simulations are also reported and confirm that the presence of a metallic-like dome at the BEC/GST interface would result in an increase of the RESET current. According to simulations, a lower dynamic resistance, coming from a better BEC/GST interface, implies a faster quench time, leading to the reported higher RESET resistance levels with cycling. So, in this case, the authors interestingly link the electrical behavior of the cell through its lifetime to the picture of elemental migration in the cell core. The same evidences have been reported for the PCM wall (heater-based) architecture at the 45 nm technology node [13] and supported by a numerical model developed by Novielli and co-workers [25]. The migration behavior has been investigated in this case after a SET pulse delivered to the cell, starting from the liquid phase and having a slow quench. Figure 3.12a

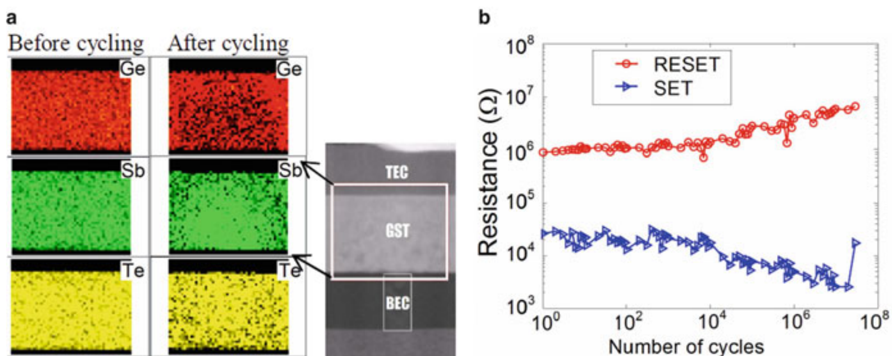


Fig. 3.11 Elemental analysis (EDX) showing the behavior of Ge–Sb–Te before and after cycling (**a**) and SET/RESET levels behavior upon cycling (**b**)

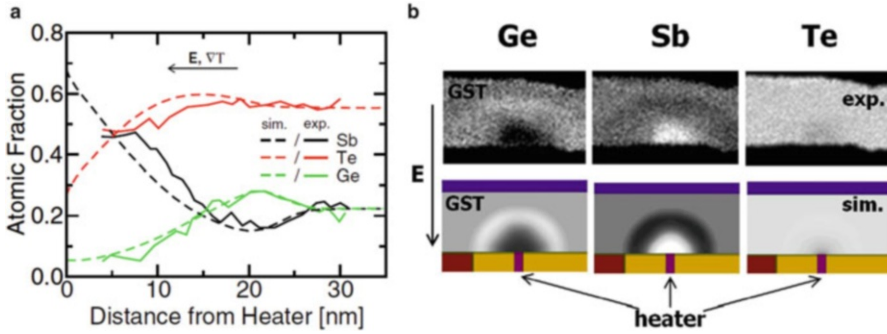


Fig. 3.12 Comparison between experimental and simulated profiles (**a**) and EDX elemental analysis and picture of the simulated migration in the *wall* architecture [13] (**b**)

shows the typical simulation results provided by the migration model described by the following equation:

$$\frac{\partial C}{\partial t} = \nabla \cdot \left[D \left(\nabla C + \frac{Z^*}{kT} C \nabla V + \alpha C \nabla \ln{(T)} \right) \right] \tag{3.6}$$

where C is the atomic concentration, D is the phase- and temperature-dependent diffusivity, Z^* is the effective charge, and α is the thermo-diffusion coefficient of each species. The expression in brackets is the atomic flux, in which the first term is the first Fick's law, the second term is the field-driven component, and the third term represents the temperature gradient component. The GST phase distribution is computed thanks to the temperature profile coming from the electro-thermal model, as will be explained in Chap. 4; then the mass transport solver leads to the space distribution of each atomic species. The computed atomic profiles can be eventually compared to experimental data from energy dispersive X-ray (EDX) analyses, as reported in Fig. 3.12a.

Qualitatively, the same space distribution is obtained, resulting into a hemi-cylindrical region closer to the heater remaining Sb-rich, while the surrounding crust becomes Ge-rich. As Rajendran et al. point out, Ge moves opposite to the electric field, E, direction due to segregation at the liquid/solid interface, modeled here as a segregation force, F_{segr}, pushing Ge atoms from the liquid to the solid phase and overcoming the E force. Owing to the slow quenching front of the electric pulse, as the liquid/solid interface moves closer and closer to the heater, more and more Ge atoms are swept away from the cell core leaving the molten region Ge-depleted, while its surrounding becomes Ge-rich (see the EDX analyses in Fig. 3.12b). The opposite trend occurs to Sb atoms: as they tend to stay in the liquid phase, this adds up with the E contribution. With the squeeze of the liquid region during ramp-down, Sb atoms concentrate near the heater electrode (BEC). Both these contributions aim at modeling the nonuniformity of the GeSbTe phase during the normal PCM device lifetime and through cycles. There are also, in the literature, other contributions that aim at modeling the same phenomena on simpler, linear, cell architectures, like the line-bridge cells. The value proposition of such devices is sustained by their very simple

architecture, enabling a very basic study of the wanted mechanisms. For example, Yang et al. report on the atomic migration in molten and crystalline $Ge_2Sb_2Te_5$ line cells, starting from the remark that when a high electric field or current is applied to materials, electric field-enhanced atomic transport, or electromigration, occurs [26]. They make a classification of the migration force into two contributions: the direct force of the electrostatic field on the diffusing ions (electrostatic force) and the momentum exchange between the moving charge carriers and the diffusing atoms (wind-force). The former happens in the molten phase only and the latter in the solid phase. A Sb depletion is reported at the anode and enrichment at the cathode, occurring after $t = 0.17$ ms stress in the molten state.

After $t = 0.53$ ms of constant biasing, the Sb and Te concentration profiles show interdiffusion of Sb and Te atoms before the diffusion of Ge atoms. Ge and Sb atoms migrate toward the cathode, whereas Te atoms migrate toward the anode. The redistribution of the $Ge_2Sb_2Te_5$ phase into the GeSb and Te phases is claimed to be due to the electrostatic force experienced by each element in the molten state. The different migration direction is possible thanks to the valence difference of each atom in the molten state: the electronegativity of each element results in Te atoms becoming anions and Ge and Sb atoms losing their electrons and becoming cations. Therefore, anionic Te diffuses to the anode, while cationic Ge and Sb diffuse to the cathode due to electrostatic forces. Also, they report that the diffusion of Sb atoms is faster than that of Ge atoms, as evaluated by computing the number of migrating atoms across the center of the line normalized by the initial concentration of each element, $\Delta C/C$, as a function of the stress time, as reported in Fig. 3.13a. From the number of migrating atoms, the authors could calculate the DZ^* product for each element, in the hypothesis that electric field and number of atoms per unit volume are constant along the line and that the temperature remains at the melting point, $T_{MELT} = 900$ K. Then electromigration was studied in the crystalline state of $Ge_2Sb_2Te_5$: Ge-rich, Sb-rich, and Te-poor regions are shown close to the anode, while Ge-poor, Sb-poor, and Te-rich regions are observed close to the cathode.

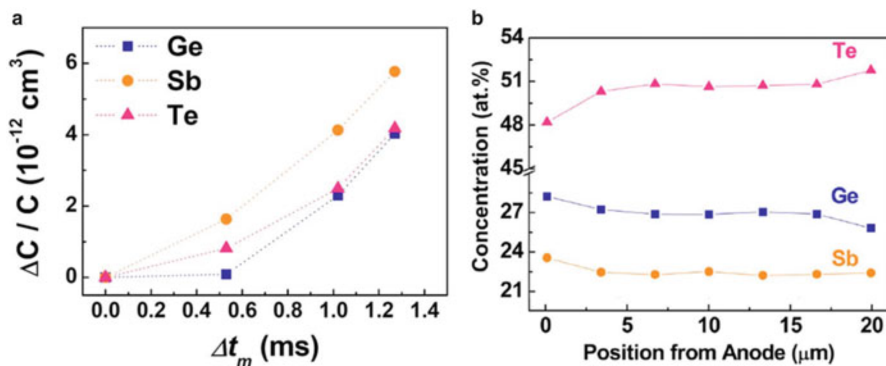


Fig. 3.13 Number of migrating atoms across the center of the line, normalized by the initial concentration of each element after an over-melting stress of duration Δt_m (**a**) and absolute concentration of elements as a function of the distance from the anode after a stress on the crystalline phase (**b**)

Despite the variation in the GeSbTe composition being negligible within the cubic phase than in the molten one, due to diffusivity variation, the concentration gradient of each element shows an opposite gradient than that shown by the molten state. The analysis reveals that the Te flux to the cathode is about twice as large as the Ge and Sb fluxes to the anode, in crystalline GST. Concerning the driving force of migration, the mass transport to the cathode comes from momentum transfer from conducting carriers, namely, wind-force, involving all the elements in the same matrix to the same direction. Generally, atoms migrate in the electron flow direction. However, the wind-force effect in crystalline GST is opposite to the well-known electron wind-force in the electromigration of metal interconnects. The concentration for each element in this case is reported in Fig. 3.13b. Nevertheless, it has been observed that for a hole-conducting material, mass transport by wind-force can be toward the cathode, which was reported for the electromigration of p-type polycrystalline silicon. Since $Ge_2Sb_2Te_5$ shows p-type transport because of intrinsic vacancies, the results by Yang et al. pave the way to a wind-force caused by hole conduction in cubic phase GST [26]. By using very similar devices, Kang et al. studied elemental separation induced by electric field, by transmission electron microscopy [27], with insights in segregation phenomena. In order to point out the cause of elemental separation in the molten phase of GST, the authors refer to the liquidus projection on the ternary phase diagram, reported in Fig. 3.14.

The GST composition is placed in the primary phase field of GeTe. Hence, in the presence of no external perturbations and by following the thermodynamic equilibrium cooling path, GST is expected to suffer from GeTe precipitation. However, under the effect of an applied electric field, tellurium is reported to be polarity

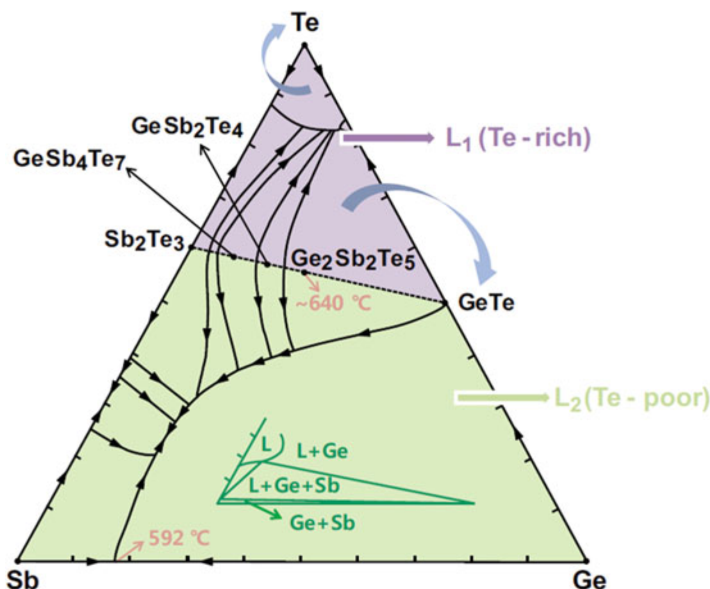


Fig. 3.14 Liquidus projection of the GeSbTe ternary diagram. Black arrows highlight a temperature decrease. The GeTe–Sb_2Te_3 pseudo-binary line represents a demarcation line between the Te-rich liquid phase (L1) and the Te-poor liquid phase (L2)

dependent, resulting into the separation of the molten phase into a Te-rich liquid and a Te-poor one. The Te-rich liquid is concentrated close to the anode, while the Te-poor one is migrated close to the cathode, as reported by other studies. The strength of the electric field and current density, responsible of such migration, has been evaluated as around 10^4 V/cm and 5 MA/cm^2, respectively, that is the typical range of PCM operation. The relative distribution of Ge and Sb is showing no field dependence in terms of direction, highlighting that their mutual separation is thermodynamic only. Despite the Ge–Sb binary system being a eutectic system, which mutually separates at solid state under the eutectic temperature, the mutual separation behavior persists when the Te content is low in the Ge/Sb enriched system. Instead, higher Te content in the system results in preferential Ge precipitation. As a general comment, the authors point out that Sb enrichment, happening at the bottom electrode, conventionally biased with negative polarity and Te enrichment, occurring at the top electrode, positively polarized, is claimed to be an intrinsic behavior of molten GST under the effect of an electric field. The segregation of Ge is explained as a lateral result of Te separation. If the degree of Te separation is high, the Ge segregation will appear surrounded by Sb. Padilla et al. introduced, for the first time, polarity experiments on PCM pore cells [28], highlighting that with a “good” polarity pulse (forward bias), Ge and Sb are pulled by bias down through the melt toward the negatively biased BEC (Fig. 3.15a). Sb might aggregate, but all local stoichiometries support rapid crystallization. Instead, with a “bad” polarity ramp-down pulse (reverse bias), Ge leaves the bottom of the pore, enabling the conditions for Sb to further segregate via crystallization at temperatures at which Te-rich material remains in the molten state, owing to

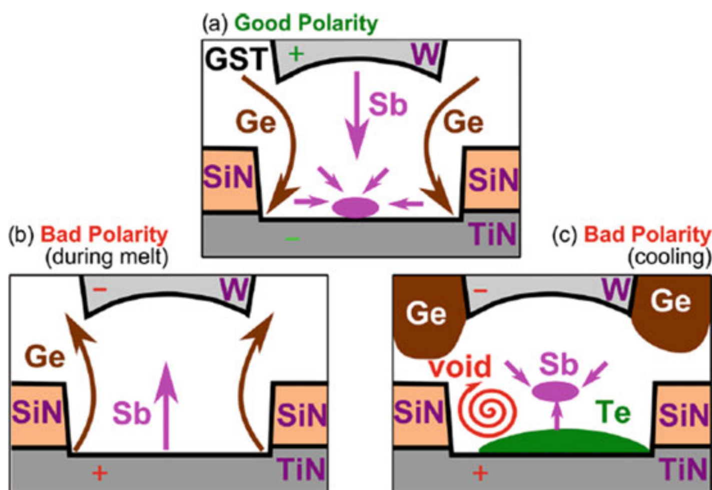


Fig. 3.15 “Good” polarity pulse, with Ge and Sb pulled down through the melt close to the cathode (**a**); “bad” polarity ramp-down pulse, with Ge leaving the bottom of the pore and Sb further segregating via crystallization at temperatures at which Te-rich material remains fluid (**b**); current focusing on a portion of the BEC, as a consequence of the Te-rich region quenching into the amorphous phase, may result into a void (**c**)

its lower melting temperature (Fig. 3.15b). If the quench of Te-rich material to amorphous focuses current into a portion of the BEC, a void can be created, resulting into reliability issues (Fig. 3.15c). Along the failure-related path, Nam et al. highlighted, by using simple multi-line structures with $Ge_2Sb_2Te_5$, that incongruent melting could be the cause of the ultimate breakdown of the lines. Indeed, the phase diagram indicates that the $Ge_2Sb_2Te_5$ compound does not congruently melt. In particular, $Ge_2Sb_2Te_5$ should be phase separated to form on one hand a Sb, Te-rich liquid phase and on the other hand a Ge-rich solid phase between 630 and 650 °C. Above 650 °C the molten state is made by a single phase, as shown in the phase diagram of Fig. 3.16b.

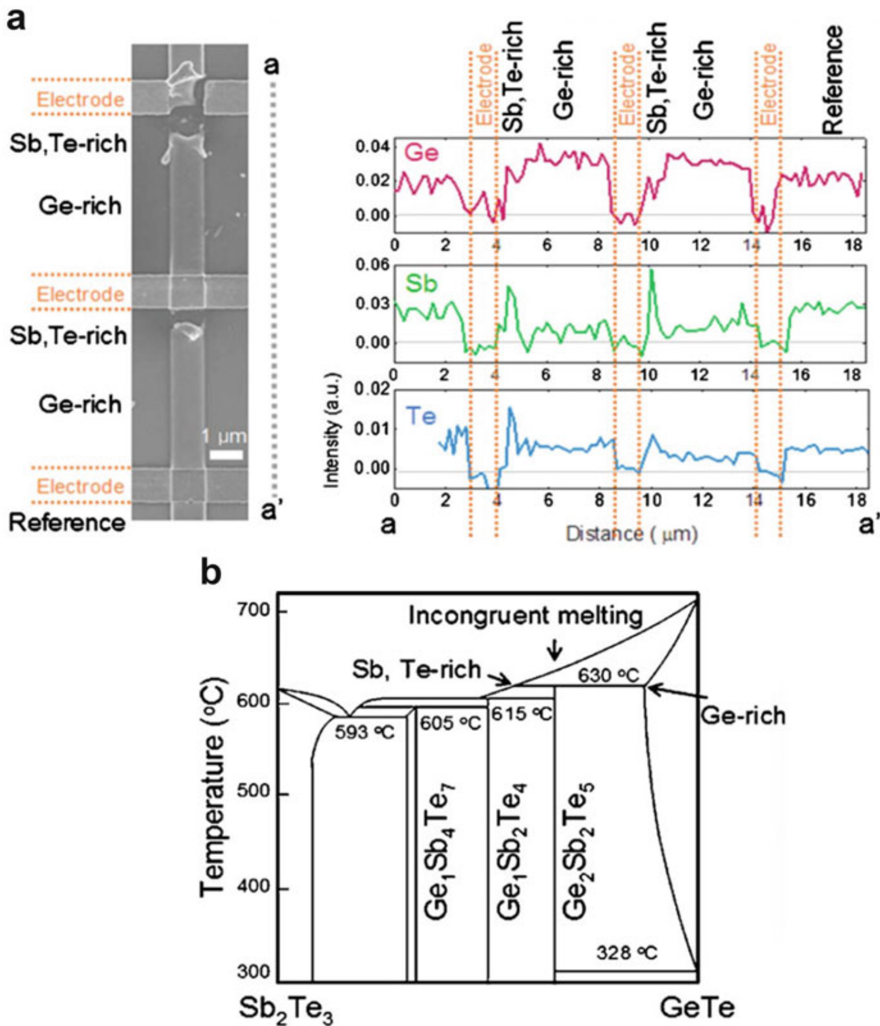


Fig. 3.16 SEM images of the structures with elemental mapping, showing a separation into two parts such as a Sb, Te-rich part and a Ge-rich part (**a**) and phase diagram along the pseudo-binary line Sb_2Te_3–GeTe, showing the incongruent melting point (**b**)

In some parts of the used line structures, temperature overcomes 630 °C and the phase separation occurs with the formation of a Sb, Te-rich liquid phase and a Ge-rich solid phase.

The accumulation of such liquid phase induces permanent failure, namely, an electromigrated portion of the line [29], reported in Fig. 3.16a. Following the same failure-oriented path, Yang et al. investigated both the migration and the void formation behavior in large and symmetric line cells with undoped and doped GST. The electromigration rate is represented by the product of the diffusion coefficient, D, and the effective charge number, Z^*, leading to DZ^*; such parameter and the amount of void formation in doped GST are quantified and compared with undoped GST. Bismuth (Bi) doping of GST increases by 10% both the DZ^* parameter and the void area, as compared with the undoped case. On the other hand, the same parameters are shown to be decreased by 30% and 45% in the O-doped and N-doped GST, respectively, as pointed out in Fig. 3.17.

According to these results, interstitial-type dopants, like oxygen (O) or nitrogen (N), could be effective in suppressing the electromigration rate and void formation by, as claimed, decreasing the free volume available for element migration; however, the substitutional-type dopant (Bi) is shown to be not effective in these terms [30]. Hence, the understanding of the atomic-bonding of dopants with the GST network can help, according to the authors, the design of phase change materials more resilient against failure. Finally, Crespi et al. propose an ion migration model, studied on the *wall* architecture [13], including mechanical stress, which is expected to play a role along with the electric field and the thermal gradient, already pointed out in Eq. (3.6). Thus, they report and solve the following continuity equation to compute the atomic migration of the three elements of GST:

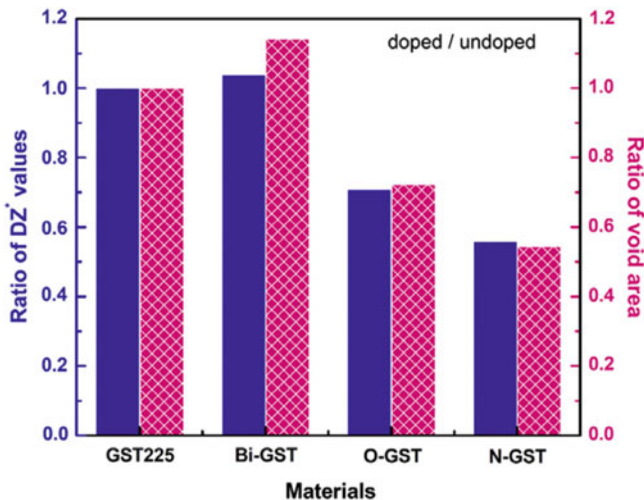


Fig. 3.17 Ratios of DZ^* values and void area, doped/undoped cases

$$\frac{\partial c_i}{\partial t} = -\nabla \cdot \left[-D_i c_i \left(\frac{\nabla c_i}{c_i} - \frac{q}{kT} z_i^* \mathbf{E} + \alpha_i \frac{\Delta T}{T} - \frac{\Omega_i}{kT} \nabla \sigma_{\text{hydro}} \right) \right] \tag{3.7}$$

In Eq. (3.7), c_i denotes the elemental concentration, D_i the corresponding diffusivity, z_i* the effective charge number, **E** the electric field, α_i the thermal diffusion factor, and Ω_i the effective atomic volume. σ_{hydro} represents the hydrostatic stress computed as the average of the x, y, and z components of the stress tensor $\boldsymbol{\sigma}$. Thermal diffusion is claimed to be the most relevant driver for both Ge and Sb. Such effect drives Sb toward the hot heater/GST interface while pushing away the Ge atoms toward the cold outer regions. The thermal coefficients of the two species are substantially the same and opposite. In both cases, the ion velocity component due to electromigration plays a marginal role, consistently with the experimental data showing no significant variation of the Ge profile by reversing bias. The role of the mechanical term is intermediate. In a real 3D simulation, where the stress along the z-axis is expected to be lower than the values computed in a 2D plain strain assumption, the mechanical term will be lower as well. Concerning the Te behavior, the most relevant driver is the electric field. In forward bias the electric field term pushes the element away from the heater electrode, while by reversing the polarity, the total ion velocity becomes negative, leading to a slight Te accumulation at the interface. Moreover, the Te transport is marginally affected by the thermal gradient [31].

3.4 Conclusions

Temperature rise, high electric field, and high current density are inherent ingredients of the phase change operation in PCM devices. The understanding and characterization of the cell thermal confinement through the study of the thermal resistance parameter and the comprehension of the temperature-related effects on the phase change material are very important steps for enabling the PCM technology. In particular, the investigation of the impact of geometry and phase change material on the cell thermal resistance is of capital importance to design cell architectures with better thermal confinement. This enables the design and control of power consumption according to the required specs dictated by applications. The study of fragility of phase change memory alloys paves the way for the research of compounds with better performance in terms of crystallization velocity and the learning on the phase change material ion migration and segregation in the molten phase as well as in the crystalline solid phase has a lot of reliability implications in terms of endurance. The research on such aspects come from very different contributors, both from the industry and the university/laboratory environments, enabling a very rich knowledge, able to boost and orient the memory makers on their development path.

References

1. S. Ovshinsky, H. Fritzsche, Amorphous semiconductors for switching, memory and imaging applications. IEEE Trans. Electron Devices. **20**(2), 91–105 (1973)
2. M. Boniardi et al., Optimization metrics for phase change memory (PCM) cell architectures. IEEE Int. Electron Device Meet. IEDM Tech. Dig., 29.1.1–29.1.4 (2014)
3. A.L. Lacaita et al., The race of phase change memories to nanoscale storage and applications. Microelectron. Eng. **109**, 351–356 (2013)
4. T. Siegrist et al., Disorder-induced localization in crystalline phase-change materials. Nat. Mater. **10**, 202–208 (2011)
5. L. Crespi et al., Electrical conductivity discontinuity at melt in phase change memory. IEEE Electron Device Lett. **35**(7), 747–749 (2014)
6. A. Redaelli et al., Threshold switching and phase transition numerical models for phase change memory simulations. J. Appl. Phys. **103**, 111101 (2008)
7. M. Boniardi et al., Electrical and thermal behavior of Tellurium poor GeSbTe compounds for phase change memory, in *Proceedings of International Memory Workshop (IMW)*, 2012
8. A. L. Lacaita et al., Electrothermal and phase-change dynamics in chalcogenide-based memories, in *IEEE International Electron Device Meeting (IEDM) Technical Digest*, 2004
9. M. Boniardi et al., Internal temperature extraction in phase-change memory cells during the reset operation. IEEE Electron Device. Lett. **33**(4), 594–596 (2012)
10. A. Sebastian et al., Crystal growth within a phase change memory cell. Nat. Commun. **5**, 4314, 1–9 (2014)
11. F. Pellizzer et al., Novel μtrench phase-change memory cell for embedded and stand-alone non-volatile memory applications, in *Symposium on VLSI Technology, Digest of Techical Papers*, 2004, pp. 18–19
12. F. Pellizzer *et al.*, A 90nm phase change memory technology for stand-alone non-volatile memory applications, in *Symposium on VLSI Technology, Digest of Technical Papers*, 2006
13. G. Servalli, A 45nm generation phase change memory technology. Int. Electron Device Meet. IEDM Tech. Dig. 113–116 (2009)
14. A. Redaelli et al., Interface engineering for thermal disturb immune phase change memory technology. IEEE Int. Electron Device Meet. IEDM Tech. Dig., 750–753 (2013)
15. B. Legendre et al., Phase diagram of the ternary system Ge-Sb-Te. I. The subternary GeTe – Sb_2Te_3 – Te. Thermochim. Acta **78**, 141–157 (1984)
16. S. Bordas et al., Phase diagram of the ternary system Ge-Sb-Te. II. The subternary Ge – GeTe – Sb_2Te_3 – Sb. Thermochim. Acta **107**, 239–265 (1986)
17. P.G. Debenedetti et al., Supercooled liquids and the glass transition. Nature Rev. Artic. **410**, 259–267 (2001)
18. C.A. Angell, Formation of glasses from liquids and biopolimers. Science **267**, 1924–1935 (1995)
19. J. Orava et al., Characterization of supercooled liquid $Ge_2Sb_2Te_5$ and its crystallization by ultrafast-heating calorimetry. Nat. Mater. **11**, 279–283 (2012)
20. J. Orava et al., Ultra-fast calorimetry study of $Ge_2Sb_2Te_5$ crystallization between dielectric layers. Appl. Phys. Lett. **101**, 091906 (2012)
21. M.D. Ediger et al., Crystal growth kinetics exibit a fragility-dependent decoupling from viscosity. J. Chem. Phys. **128**, 034709 (2008)
22. G.C. Sosso et al., Dynamical heterogeneity in the supercooled liquid state of the phase change material GeTe. J. Phys. Chem. B **118**, 13621 (2014)
23. B. Rajendran et al., On the dynamic resistance and reliability of phase change memory. Symp. VLSI Technol. Dig. Tech. Pap. 96–97 (2008)
24. M. Boniardi et al., Study of cycling-induced parameter variations in phase change memory cells. IEEE Electron Device Lett. **34**(7), 882–884 (2013)
25. G. Novielli et al., Atomic migration in phase change materials. IEEE Int. Electron Device Meet. IEDM Tech. Dig. 589–592 (2013)

26. T.-Y. Yang et al., Atomic migration in molten and crystalline $Ge_2Sb_2Te_5$ under high electric field. Appl. Phys. Lett. **95**, 032104 (2009)
27. D. Kang et al., Analysis of the electric field induced elemental separation of $Ge_2Sb_2Te_5$ by transmission electron microscopy. Appl. Phys. Lett. **95**, 011904 (2009)
28. A. Padilla et al., Voltage polarity effects in $Ge_2Sb_2Te_5$-based phase change memory devices. J. Appl. Phys. **110**, 054501 (2011)
29. S.-W. Nam et al., Phase separation behavior of $Ge_2Sb_2Te_5$ line structure during electrical stress biasing. Appl. Phys. Lett. **92**, 111913 (2008)
30. T.–.Y. Yang et al., Influence of dopants on atomic migration and void formation in molten $Ge_2Sb_2Te_5$ under high-amplitude electrical pulse. Acta Mater. **60**, 2021–2030 (2012)
31. L. Crespi et al., Modeling of Atomic Migration Phenomena in Phase Change Memory Devices, in *Proceedings of International Memory Workshop (IMW)*, 2015

Chapter 4
Self-Consistent Numerical Model

Andrea Redaelli

4.1 Introduction

As extensively discussed in the previous chapters, the ternary compound $Ge_2Sb_2Te_5$, (in the following GST) is the reference alloy employed in the realization of phase change memories (PCMs). The high conductive GST state is related to a polycrystalline phase organized in an octahedral environment (NaCl-like) where one sub-lattice is fully constituted by tellurium atoms (sixfold coordinated), and the other one is randomly occupied by germanium and antimony [7, 50, 61, 80, 84]. The germanium and antimony sub-lattice presents one atomic position unoccupied for each basic cell due to the stoichiometry, resulting in 20% vacant sites [59]. A long-range structural order is thus present in the crystalline phase, so it can be described in the framework of a Bloch semiconductor where vacancies act as dopants, pinning the Fermi level close to the valence band [68]. Concerning the low conductive state, the absorption spectra in amorphous GST show also the existence of an optical gap of about 0.7 eV with an exponential Urbach tail with a decay of 80 meV, related to localized states (see also Chap. 2). The existence of localized states in the bandgap is consistent with the disordered nature of the amorphous phase, as also reported by ab initio simulations for similar systems [70]. Since the transitions induced by programming operations occur with very fast times (10–200 ns) and since more than 10^9 switching operations have been demonstrated on prototype devices [71], it is reasonable to assume that GST is very flexible (very high atomic diffusivity till temperature down to the glass transition as demonstrated in GeTe system by molecular dynamics in [33]) and the bonding between the elements can change easily due to the high mobility of atoms in the supercooled molten material as well as in overheated amorphous. This picture is consistent with

A. Redaelli (✉)
Micron Semiconductor Italia Srl, Vimercate, Italy
e-mail: aredael@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_4

GST peculiar fragile behavior as extensively discussed in Chap. 3. From ab initio simulation, a special role has been highlighted for germanium which is in the octahedral coordination in the octahedral crystal phase, while coexistence of defective octahedral, pyramidal, and tetrahedral coordination is present in the amorphous phase [22, 45]. From atomistic simulations in similar system (GeTe), Ge-Ge homopolar chains have been highlighted as key catalyzer for the transition between amorphous and crystalline phases promoting enhanced mobility and fast crystallization speed through the so-called dynamical heterogeneity [35]. More details about the structure of crystal and amorphous phases are reported in Chap. 6.

Despite the physics of chalcogenide glasses at microscopic level is very complex and often requires atomistic approaches, in this chapter we will focus on simplified models that are, in the last section, coupled in a fully consistent framework able to describe the macroscopic features of the PCM devices with a finite element approach. The main focus will be given to the threshold switching of the amorphous phase, that is, an element of novelty with respect to conventional electronics as well as to the crystallization kinetics.

4.2 Threshold Switching

Threshold switching (TS) is a fundamental mechanism for PCM devices, enabling a high current density in the device active area of amorphous bits, without developing enormous electrical fields. Despite this fundamental role of the threshold switching, the nature of the physical principle responsible for the conductance switching is still controversial [3, 7, 23, 26, 27, 30, 31, 36, 39, 55, 58, 67, 78, 85]. In the past 50 years, several theoretical efforts have been devoted to explain this crucial phenomenon. Due to the large thermal activation energy of amorphous chalcogenide electrical conductivity (see Chap. 2), the switching phenomenon was first ascribed to a thermal effect. By increasing the applied voltage, in fact, the current density increases, and, as a result of Joule effect, the temperature increases too, thus locally enhancing the electrical conductivity. When the temperature sufficiently rises, a positive feedback switches on, and a huge thermal generation of electrical carriers takes place. The voltage drop on the amorphous film can thus decrease while maintaining a high current density [2, 81]. In 1978, Adler et al. proposed that the effect is not thermal in nature, but, in agreement with the original Ovshinsky hypothesis, he explained the TS as a pure electronic phenomenon [26]. In fact, he supported a purely electronic mechanism primary related to the balance between a strong Shockley-Hall-Read (SHR) recombination through trap levels and a generation mechanism driven by both electric field and carrier densities [27]. This approach has been analytically and numerically modeled, and it has been proven as a possible way to simulate the switching behavior by using a set of simple equations available in standard simulation tools [13]. More recently, alternative works provided different points of view. Karpov and coworkers [55, 78] proposed a model based on interplay between the electronic transport and the nucleation

kinetics. They preserved the original idea of conductive filaments as responsible for the switching behavior. Starting from an energetic balance, they get to a model based on crystallization of a filament that expands in the amorphous matrix until shorting the surrounding amorphous region by connecting the electrodes. Ielmini and coworkers [39, 85] instead proposed a switching model consistent with the Poole-Frenkel transport model, that is, the thermally assisted trap-limited conduction extensively discussed in Chap. 2. In this framework, the switching is ascribed to the increase in the average kinetic energy of the carriers through the length of the device, resulting from the balance between the field-induced energy gain and the energy relaxation in the lattice due to the scattering with phonons. This model leads to a nonuniformity of the electric field, similarly to the one reported in [13]. In the original works, the proposed model (both in the amorphous transport and switching parts) was not self-consistently proven by coupling the proposed transport equation with the drift-diffusion and the Poisson equations. Piccinini and coworkers [23, 30, 58] developed a Monte Carlo model that proved the self-consistency between drift-diffusion equation, Poisson equation, and thermally assisted trap-limited conduction. Today, we can say, with a certain degree of confidence, that the Poole-Frenkel mechanism is the main ingredient to explain amorphous chalcogenide electronic transport. However, what happens close to the threshold point is still not completely clear. Evidences of avalanche multiplication have been reported in this regime by Paolo Fantini [65] as key element to explain the experimentally observed superlinear temperature behavior of the IV characteristics close to the switching point.

In this section of the chapter, we will focus on the description of the original Adler model as it is reviewed and discussed in [13]. The model is very general and is based on a balance between a carrier recombination mechanism with a carrier generation mechanism. This model can be simply embedded in a commercial simulation tool with conventional equations employed for semiconductor description, till maintaining a good physical sound in respect to amorphous transport physics.

In his original work, Adler considered the chalcogenide material as a doped homogeneous semiconductor resistor with a single trap level close to one band and a doping level close to the opposite band (e.g., an acceptor doping density N_A close to the valence band and an N_D donor trap concentration near the conduction band). This description represents an attempt to describe the electrical behavior of chalcogenide glasses with a minimum number of assumptions. The acceptor doping density N_A is assumed greater than N_D, consistently with the p-type low-field conductivity experimentally observed on majority of chalcogenide compounds [15, 69]. Note that, although the original model [27] relies on the microscopic description of the chalcogenide compound based on valence-alternation pairs hypothesis [47], its validity can be extended to any system characterized by defects states that act accordingly (in this case as donor-like traps).

Since the N_A and N_D are of the same order of magnitude, at thermal equilibrium, the Fermi level of the system is expected to be placed close to the middle of the gap, and the density of ionized traps $N_D{}^+$ practically equals the total trap concentration

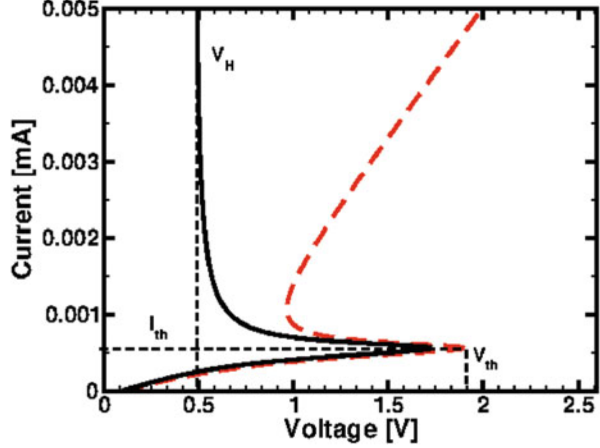


Fig. 4.1 The solid curve shows the current-voltage curve obtained with the analytical model considering only the chalcogenide material. When the presence of a contact resistance is also considered, the *dashed curve* is obtained (Reprinted with permission from [13]; copyright AIP Publishing LLC 2008)

N_D. Applying an external voltage to the system, the free carrier number should increase due to generation and the traps start to react with the electrons thus, partially filling and neutralizing the N_D^+ centers. The probability for a positive charge trap to capture a hole and for a neutral trap to capture an electron will be neglected in the following. In steady state, the electrostatic potential Φ, the electron concentration n, and the hole concentrations p can be computed by solving the coupled set of Poisson and current continuity equations:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot J_n + G_n - R_n \tag{4.1}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla \cdot J_p + G_p - R_p \tag{4.2}$$

$$\nabla \cdot \epsilon \nabla \Phi = -\rho(\Phi) \tag{4.3}$$

In Adler's system, the main recombination mechanism is the SHR through trap levels, while a field-dependent carrier-assisted generation rate is considered. The carrier generation is mainly ascribed to a mechanism that is a monotonic function of both the electric field and the carrier densities, like the avalanche multiplication. Detailed calculations are reported in [13].

Figure 4.1 shows the resulting current-voltage (I-V) curve for a resistor of length 10 nm and square section 1 × 1 μm^2. The solid curve represents the I-V characteristic obtained from the analytical model, thus related to the chalcogenide resistor alone. The dashed one has been obtained including a series resistance, usually related to the external contacts. As expected, both curves present an off-state region where the current increases with increasing voltage. When a threshold voltage V_{th} is reached, the I-V curves snap back to the on-state region. Note that in the on condition, the voltage drop across the chalcogenide remains constant. Practically, the chalcogenide in the on state behaves as a constant voltage generator. The holding voltage V_h is thus the voltage drop on the chalcogenide alloy in the highly

conductive state, and the resistance in the high conductive state R_{on} is practically the contact one. The obtained result is qualitatively consistent with the observed experimental evidences. It is worth to stress once again that this effect is a crucial feature of chalcogenide material that allows a high current flowing in the device thus promoting the temperature required for the phase transition through Joule heating. The next section will focus on the modeling of the memory switching, that is, the permanent transition from the amorphous to the crystalline phases after a proper heating of the material.

4.3 Memory Switching

The threshold switching mechanism, descripted in the previous paragraph, refers to a reversible switch between a low conductive to a high conductive state as a function of the applied bias. When the applied voltage exceeds the threshold voltage of the chalcogenide material, the current starts flowing in the device, but, if the chalcogenide is not maintained in a biased condition, the low conductive state is spontaneously recovered. Threshold switching itself cannot thus be employable for nonvolatile memory cells. However, among the materials that feature threshold switching, some compounds, such as GST, can permanently store the high conductive state also without applied bias. For these compounds, once high current starts flowing, the chalcogenide material heats up above its crystallization temperature, and the phase change can eventually occur (MS, memory switching). Since the PCM cell performances are strongly related to the phase transition kinetics, understanding of MS physics represents a key issue for the development of this technology. It is worth noting that the threshold switching transition is electronic in nature and distinct from the memory effect for chalcogenide materials showing the phase change transition. This has been experimentally demonstrated in PCM devices, where the threshold switching can take place without a transition from the amorphous to the crystalline phase and where, despite the device switches in the on state, higher currents (or longer times) are required to achieve the memory phase change [10].

4.3.1 Memory Switching for Optical Data Storage

The description of the phase change in chalcogenide materials is strictly related to the physics of the phase transitions in solid systems. These theories were originally developed in the late 1930s and 1940s [16, 17] for metals, but several efforts have been made to exploit them for the investigation of the phase change transitions in chalcogenide compounds, especially for optical data storage applications [41, 42, 44, 52, 59, 61, 79, 83]. Main focus of those works was the research of new materials with enhanced performances in terms of optical contrast, high transition speed,

Fig. 4.2 Ternary phase diagram showing the materials employed in optical data storage and related years of discovery (Reprinted with permission from [57]; copyright Nature Publishing Group 2007)

good thermal stability to ensure data retention, good capability to sustain a large number of transition, as well as a good chemical stability. In the mid-1980s, GeTe material was found to show both fast crystallization and good contrast [52], thus triggering the interest in the GeTe-based compounds, such as the ones belonging to the GeTe–Sb_2Te_3 tie line [41, 42, 61] which were employed in commercial products [43, 46, 59, 73, 79]. Subsequently, another chalcogenide family based on the $Sb_{70}Te_{30}$ eutectic alloy and generally indicated as doped SbTe compounds has been studied and employed [37, 43, 51, 54] including doping with In (e.g., $In_x(Sb_{70}Te_{30})_{1-x}$) and doping with Ag and In (e.g., $Ag_xIn_y(Sb_{70}Te_{30})_{1-x-y}$), also called "AIST."

In Fig. 4.2, the ternary phase diagram for materials employed in optical data storage is reported showing the compounds used as well as the year of discovery for each disk generation as reported by Wuttig and Yamada in Ref. [57]. Research in the optical storage field in the past 30 years allowed the development of a lot of knowledge useful also in the solid-state memory field. In addition to the optical disk requirements, chalcogenide materials for electronic devices must have a good electrical contrast, i.e., difference in electrical conductivity of the two phases. Furthermore, phase transition in a solid-state device presents peculiar features such as the volume constraints and interfaces with other materials. In the following sections, the physics that describes the phase transformations in chalcogenide compounds is described, focusing on the two main phase transitions: the crystal-to-amorphous transition (vitrification) and the amorphous-to-crystal one (crystallization).

4.3.2 Glass Transition

Glass transition (or vitrification) is defined as the transformation process from a liquid phase to a solid glass. In particular, the solid that results from the cooling of the liquid strongly depends upon the cooling rate. At low cooling rates (below some critical value), the liquid solidifies creating a crystalline phase. Instead, if the cooling is faster, the alloy freezes without crystalizing, becoming an undercooled liquid and then transforming into a solid glass. To realize the transition from the crystalline state to the amorphous one, GST has thus to be kept above the melting point (620 °C) such that transformation into liquid phase firstly occurs and then rapidly quenched to avoid crystallization during cooling. The result is thus a frozen liquid phase that solidifies into a solid disordered status [29]. Typical cooling rate required for the GST alloy is about 10^{10}–10^{11} K/s [29]. However, demarcation line between cooling rates that provides a glass or a crystal is not always clear. The most common metric used is the so-called glass transition, Tg. Tg is empirically defined as the temperature corresponding to the appearance of discontinuities of certain important material properties such as specific volume, thermal expansion coefficient, specific heat, and viscosity. Above Tg an undercooled liquid is able to maintain a state of internal equilibrium. Its free energy is, in any case, higher than that of a crystalline solid, resulting in metastability. However, due to the negligible atomic diffusion, no internal equilibrium can be achieved below Tg, thus preventing, in a certain range of temperature and time, crystallization of the metastable amorphous phase. Although transition to the amorphous phase is at the basis of the programming operation of a PCM device, it is worth noting that the description of the vitrification process in memory devices has been mainly limited to the characterization of a few phenomenological quantities, typically the current required to achieve the melting temperature. The lack of a detailed analysis is partially justified by the fact that, in a PCM device, the time required for melting and the quenching time are mainly limited by the delay between the application of the external electrical pulse and the temperature increase in the active area of the cell. Since the volume involved in the phase transition of nonvolatile memory device is in the range of 10^3–10^5 nm^3, the thermal delay results in few nanoseconds, a value that is usually larger than the vitrification time (less than 1 ns, Refs. [20, 72]). Furthermore, the typical thermal constants of the device (few nanoseconds) are in general faster and completely negligible with respect to the typical memory access time currently required in PCM applications (15–150 ns). On the other hand, the transition from a solid crystalline phase to a melted one usually occurs with absorption of a certain amount of energy that depends on the volume involved in the transition and on the latent heat of fusion. Once the temperature in the material reaches the phase transition, the minimum energy required could limit the kinetics of the melting event by imposing a simple reciprocal law between the power released to the material and the time required for the transition. However, up to now, there is not clear evidence that this effect has a major impact on the PCM device performances. There is, instead, another effect currently neglected but with

potential large implications on the functionality of PCM devices. The transition from liquid to undercooled liquid to a glass is usually related to a decrease in the specific volume [29]. It is worth noting that the specific volume of the glass is, in any case, greater than that of the crystal at all temperatures where the two phases may coexist, due to the so-called free volume that is related to the more open structure in the glass, usually few percent (6% for GST, Refs. [4, 7]). Since the GST vitrification results in the transition to a less dense disordered phase and the amorphous region in a PCM cell is confined by the crystalline background, GST vitrifies under a hydrostatic pressure that can be in the order of several gigapascals [4]. However, this effect has been only recently addressed on bulk GST film, but an analysis of a PCM cell is still lacking due to the limited capability to experimentally access the thermodynamical variables of the system.

4.3.3 Crystallization

The crystallization of an amorphous sample requires a certain thermal budget to occur, i.e., a certain temperature for a certain amount of time. It proceeds from nucleation of small crystallites and their subsequent growth. The theory of crystallization kinetics was largely developed by Avrami at the end of the 1930s, and it can be used to describe the temporal behavior of phase changes [1, 16–18, 49, 59–61]. Starting from nucleation and growth rates, the Avrami theory under the hypothesis of isothermal condition (all the material at the same temperature) and in isokinetic regime (same temperature dependence of both nucleation and growth processes) leads to an analytical integral expression (JMA model) that relates time and temperature to the total transformed fraction:

$$x(t) = 1 - e^{-A(T)t^n} \tag{4.4}$$

where $x(t)$ is the total transformed fraction, n is the Avrami coefficient, and $A(T)$ is a temperature-dependent effective rate. Such rate can be expressed as

$$\mathrm{A(T)} = ve^{-E_a/(K_B T)} \tag{4.5}$$

where E_a is the activation energy of the crystallization and υ is the frequency factor. This equation is useful to describe the physical parameters that linearly depend on the crystal fraction, e.g., the reflectivity, and the specific heat. Although this approach can be used to model the behavior of optical data storage (reflectance is proportional to the crystal fraction), the JMA equation cannot be directly used in electronic device memory modeling. In fact, due to the existence of percolative effects [11, 48], the relationship between crystal fraction and electrical conductivity is not linear. Let us assume a homogeneous nucleation of conductive crystallites in the resistive amorphous layer sandwiched between the top and the bottom electrodes. Conduction can take place only if a crystal path appears between the two

electrodes, requiring a minimum amount of crystal to strongly increase the conductivity of the system (percolative threshold) [14, 76]. Further nucleation has then minor impact on the overall conductivity. Moreover, in the programming operation regime, both the hypotheses of isokinetics and isothermal regime are not satisfied. It follows that the integral model proposed by Avrami cannot be applied to the investigation of the crystallization kinetics in PCM devices, but a local description that considers the time and temperature dependence for both nucleation and growth is mandatory. Few numerical models that consider a local nucleation and growth algorithm have been so far proposed for optical data storage. Peng et al. [21] proposed a local model based on the standard description of quasi-steady-state homogeneous nucleation and growth, tailoring the thermodynamic parameters on experimental data. A more complex description has been proposed by Senkader and Wright [82] that considers a transient model based on rate equations, including heterogeneous nucleation. Although this model seems to be the most physically based one, it shows typical crystallization times quite different from the values measured in PCM devices [5].

4.3.3.1 Nucleation

To model the crystallization phenomenon, some thermodynamic aspects of phase transition classical theory have to be reviewed. Consider the generic transition from a phase l to a phase x at a temperature T lower than transition temperature T_m and consider, from an energetic point of view, the situation of a creation of an x sphere inside the l matrix. The free energy of a sphere may be easily expressed as:

$$\Delta G_{\text{sphere}}^{l \to x} = 4\pi r^2 \gamma + \frac{4}{3}\pi r^3 \Delta g \tag{4.6}$$

where γ is the surface energy, r is the radius of the sphere, and Δg is the bulk free energy gain corresponding to the undercooling $T - T_m$. The first term is related to the work required to build the sphere surface, while the second one is related to the bulk free energetic gain related to the undercooling. Obviously, for $T = T_m$, there is no bulk energetic gain in the phase transition, and the sphere creation does not spontaneously occur. For $T < T_m$, the bulk factor becomes negative, thus realizing the possibility of obtaining stable spheres for certain radius values. The stability of the sphere is related to the capability of adsorbing new atoms from the l phase and thus is related to the derivative of Eq. (4.6) with respect to r. When the derivative is negative, the transition from the sphere of radius r to the sphere with radius $r + dr$ is possible due to the resulting free energy gain. Imposing such condition, a critical radius may be obtained as:

$$r_c = \frac{2\gamma}{\Delta g}. \tag{4.7}$$

It is worth noting that the critical radius is infinity at the melting point when the bulk energy gain is basically zero. It instead decreases for decreasing temperature, when the energy gain moving from the supercooled liquid to the solid phase is becoming large. Starting from this simple point, it is possible to compute the energy of the critical nuclei at a given temperature and the density of critical nuclei, available for the transformation, as a function of the temperature (detailed calculations can be found in [13]. The result is reported here:

$$C_n = \rho e^{-\left(\frac{\Delta G_C}{k_B T}\right)} = \rho e^{-\left(\frac{-16\pi\gamma^3}{3\Delta g^2 k_B T}\right)} \tag{4.8}$$

where ρ is the cluster density. Once an *l* atom is attached to the critical cluster, crystallization occurs and the resulting nucleus is stable. The rate of nucleation (number of nucleation for volume and time unit) may be thus easily expressed as:

$$\mathrm{I} = C_n \frac{dn}{dt}. \tag{4.9}$$

where dn/dt is the rate of l atoms (liquid) joining a critical cluster through its surface and

$$\frac{dn}{dt} = \nu s_c p e^{\left(-\frac{E_{an}}{k_B T}\right)} \tag{4.10}$$

where E_{an} is the activation barrier at the surface, ν is the vibration frequency of an l atoms, s_c is the number of l atoms facing the x cluster, and p is the probability to jump in the cluster. Such a probability may be expressed as the product $p = f \cdot A$ with f the probability that an l atom is vibrating toward the nucleus and A the probability that it does not bounce back by an elastic collision with an x nucleus. Assuming that, during the whole transformation, the critical cluster concentration is always equal to the equilibrium one (steady-state approximation), the rate of nucleation may be expressed combining previous Eqs. (4.8) and (4.10) in Eq. (4.9), and assuming that the bulk energy gain is at the first order equal to $\Delta g = \delta(T - T_m)$, it results in:

$$I = K e^{\left(-\frac{E_{an}}{k_B T}\right)} e^{-\left[16\pi\gamma^3/3[\delta(T-T_m)]^2 k_B T\right]} \tag{4.11}$$

where $K = \nu\, s_c\, p\, \rho$ summarizes all the constants, evaluated for a nucleus at critical size. The first factor, related to the neighbor atomic diffusion, is temperature activated with an activation barrier E_{an}, while the second one, related to the existence of critical nuclei for a given undercooling, is deactivated approaching the transition temperature. For low temperatures, there are many critical clusters, but no atom will diffuse into a nucleus. On the other hand, for temperature closer to the transition, many atoms could diffuse into a nucleus but no critical clusters are available. The simple calculations proposed underline the main temperature dependencies of the nucleation mechanism. A more detailed description can be found in

the work of Turnbull and coworkers [24, 32, 40]. It is worth noting that the nucleation rate of Eq. (4.11) does not show any time dependence and it must be considered a steady-state nucleation rate. However, it has been reported in the literature that transient effects during nucleation may have an impact in the early stage of the phase transition mechanism [24]. Moreover, the simplified expression (4.11) does not consider any effect of surfaces and boundary materials on nucleation, thus neglecting the possible impact of heterogeneous nucleation reported in the literature for several materials [24]. The assumptions of the crystallization model and on the impact of the approximations in Eq. (4.11) will be further discussed in next section.

4.3.3.2 Growth

Once a critical cluster has been transformed into a stable crystalline nucleus, more atoms can diffuse into the nucleus thus increasing the size of the crystallite. To evaluate the growth rate of a nucleus (or at an interface with a previous existing crystal phase), an interface between the l and x phases has to be considered, where E_{ag} is the diffusing energetic barrier and dg is the free energy gain in attaching an atom of l to an x nucleus. For such a system, the net rate considering the jump rate of atoms from l to x minus the jump rate from x to l may be expressed as:

$$R = S\nu_0 e^{\left(-\frac{E_{ag}}{k_B T}\right)}\left[1 - e^{-\left(\frac{dg}{k_B T}\right)}\right] \tag{4.12}$$

where ν_0 is the atomic vibration frequency and S is the number of atoms at the interface. Detailed calculations can be found in [13, 16]. The growth velocity can thus be evaluated as the product of the jumping atom rate for the average jumping distance, λ, assumed equal to the lattice parameter of x phase, for unit area:

$$V = \beta\frac{\lambda R}{S} = \beta\lambda\nu_0 e^{\left(-\frac{E_{ag}}{k_B T}\right)}\left[1 - e^{-\left(\frac{dg}{k_B T}\right)}\right] \tag{4.13}$$

where β is a geometrical factor, related to the growth mode of the interface (the previous simplified calculation referred to a mono-dimensional interface). Since dg increases with the undercooling, the temperature dependence of the second factor is practically negligible over a wide range of temperatures. At temperatures close to the transition (where $dg = \Delta(T - T_m) \rightarrow 0$), the term in the brackets goes to zero, thus showing that growth velocity is 0 at T_m and it increases linearly with the undercooling near transition temperature. This is easily explained by considering the fluxes: at the transition temperature, both the direct and inverse barriers have the same heights, thus equalizing the direct and inverse atom fluxes through the interface.

4.4 The Numerical Model

The previously described models for electrothermal transport behavior have been implemented in a semiconductor solver where an external module accounts for glass transition, nucleation, and growth dynamics. In this section the numerical model and its implementation are reported.

4.4.1 Electrical Transport Model

The physics of both crystalline and amorphous phases has been mapped in a standard semiconductor model by properly implementing band diagrams and transport properties. Absorption measurements on crystalline GST show an optical indirect transition with an energy gap of about 0.5 eV. Since the conduction is mainly related to free holes with a measured density of about 10^{19}–10^{20} cm^{-3}, and very low dependence on temperature [7, 19], cGST has been modeled as a quasi-degenerate semiconductor. Because no measurement of the density of states has been documented in the literature, we assume the values reported in Table 4.1, in rough accordance with previously referred values for other chalcogenide alloys [53]. Since the density of vacant sites is about 20% of the total atomic density and at thermal equilibrium they are negatively charged (acting as a conduction band), an acceptor level with a density of 5×10^{20} cm^{-3} has been introduced in the band diagram, as reported in Fig. 4.3 on the left.

To take into account the low temperature dependence of conductivity, the trap level has been positioned at 100 meV from the valence band. As a result of the Poisson equation, the Fermi level lies close to the valence band (at about 50 meV), thus causing the *p*-type quasi-degenerate behavior of crystal (Fig. 4.3 on the left). As previously discussed, the amorphous phase can also be correctly described with a band diagram. The absorption spectra in the amorphous phase show the existence

Table 4.1 Numerical values of the parameters describing the band structure for crystalline and amorphous GST

Property	Crystalline GST	Amorphous GST
E_{gap} (eV)	0.5	0.7
N_c (cm^{-3})	2.5×10^{19}	2.5×10^{19}
N_v (cm^{-3})	2.5×10^{19}	10^{20}
Vacancies (cm^{-3})	5×10^{20}	NA
Lone pairs (cm^{-3} eV^{-1})	NA	10^{21}
Donor traps(cm^{-3})	NA	10^{17}–10^{20}
Acceptor traps (cm^{-3})	NA	10^{17}–10^{20}

Adapted with permission from [7]; copyright IEEE 2004

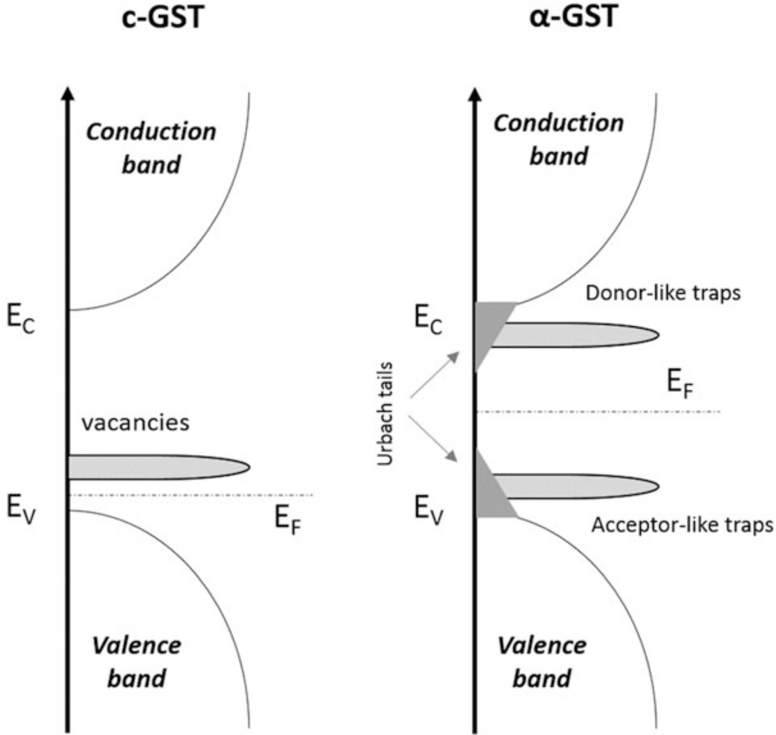


Fig. 4.3 Band diagrams for both crystal and amorphous phases implemented in the numerical model (Adapted with permission from [7]; copyright IEEE 2004). Note the huge amount of low mobility state required for the amorphous GST modeling

of an optical bandgap of 0.7 eV [7, 19]. A strong exponential activation of the electrical conductivity with temperature has been reported in the literature, suggesting that amorphous GST behaves practically as an intrinsic semiconductor with a Fermi level at about 300 meV from the free hole band [8, 19]. Furthermore, absorption spectra also show the existence of an Urbach tail with a decay constant of about 80 meV, thus extending for about 200 meV inside the bandgap [19]. To map the Mott and Davis mobility model in our model [13], we translate the amorphous band diagram with mobility tails (reported also in Chap. 2) into the band diagram of Fig. 4.3 (right) with localized state at fixed energy just below and above the conduction and valence bands. The low mobility tail, underlined by the Urbach tail and possibly related to the lone pairs, is a signature of the existence of a mobility edge in the valence band. As depicted in Fig. 4.3, we model this tail with donor states, close to the free hole band, which are neutral at thermal equilibrium with a uniform density of 10^{20}–10^{21} cm^{-3} eV^{-1} in an energy range of 200 meV from the valence band mobility edge. On the other hand, VAPs have been modeled as couples of donor-acceptor levels with a density of 10^{18}–10^{20} cm^{-3}, respectively, close to the conduction and valence bands. A drift-diffusion electrical transport

model has been implemented for both phases. For the hole mobility of the crystal phase, a value of 23.5 cm^2/Vs has been assumed leading to an overall *c*GST resistivity of about 17 mΩcm in accordance with the value reported in the literature for the crystalline fcc phase [19, 38]. Since the *n*-type conduction has no effect on transport in *c*GST and since no electron mobility data are available in literature, we assumed a negligible electron mobility of 0.1 cm^2/Vs (more than two orders of magnitude lower than hole mobility). By assuming the band diagram previously described for the amorphous phase, a mobility in the range of 150–200 cm^2/V s is required to fit the melt-quenched amorphous resistivity value of 25–30 Ω cm. It is worth noting that the large value of the assumed hole mobility can be mainly ascribed to the simplified assumption of the numerical model, where the energy-dependent hole mobility has been approximated with a digital transition between zero-mobility trapping states and a delocalized band with a nonphysical large constant mobility. A more precise description is provided in Chap. 2. The proposed hole mobility values must be thus retained as "effective" mobility values that account for the observed electrical resistivities. Also for the *a*GST, since no data are available for electron mobility and since it has no influence on the transport behavior, we assumed a negligible electron mobility of 1–10 cm^2/V s. Because of high trap level densities in both phases, the Shockley-Read-Hall model for recombination has been introduced in the model. The cross section of all the traps have been supposed to be 10^{-12} cm^{-2} for the charged traps and 10^{-16} cm^{-2} for the neutral traps.

The avalanche Okuto-Crowell model has been implemented for both phases, with a critical field of 3.3×10^5 V/cm [34]. Among the different avalanche models reported in literature, the Okuto-Crowell model has been chosen only for convergence issues. Low-field mobilities and avalanche critical fields for both phases are reported in Table 4.2. In Fig. 4.4, the low-field *I*-*V* curve of the off state of the amorphous GST is reported. Up to about 200 mV, a linear behavior is found, while for higher voltages, the current increases exponentially up to the threshold switching. To explain both the low-field behavior and the threshold switching effect, a field-assisted generation mechanism sustained by the free carrier concentration (e.g., avalanche ionization) has been considered as previously described. The competitive interaction between this generation mechanism and a strong SHR recombination caused by the large defect density is responsible for the switching behavior.

The calibrated semiconductor-like model is able to provide a good agreement with measurements, as shown, for instance, in Fig. 4.4. Figure 4.5 shows that the

Table 4.2 Numerical values for the electronic transport in amorphous and crystalline GST

Property	μ_n (cm^2 V^{-1} s^{-1})	μ_p (cm^2 V^{-1} s^{-1})	F_c (V cm^{-1})
Crystalline GST	0.1	23.5	3.3×10^5
Amorphous GST	5	200	3.3×10^5

Adapted with permission from [7]; copyright IEEE 2004

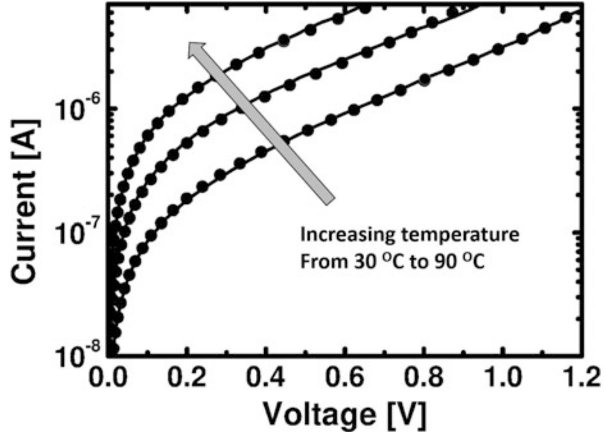


Fig. 4.4 Off-state current-voltage electrical characteristics of the amorphous phase for different external temperatures. Note the good agreement between experimental data (*black circles*) and the simulation results (*solid lines*) obtained with the equivalent semiconductor-like model for the amorphous GST (Reprinted with permission from [12]; copyright IEEE 2005)

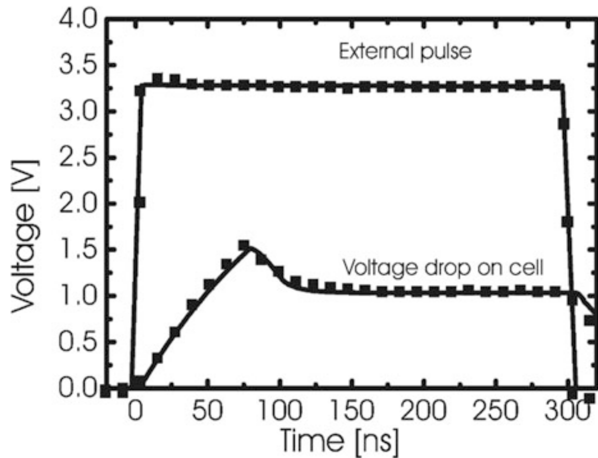


Fig. 4.5 Comparison between a simulation results (*continuous line*) and experimental data (*black squares*). Applying an external voltage pulse, the model is able to reproduce the voltage drop on the PCM device, correctly describing the switching behavior of amorphous phase also in transient condition (Adapted with permission from [12]; copyright IEEE 2005)

numerical model correctly describes the pulsed operation mode of PCM cells. The result was obtained by simulating the external electrical net present in the real experiment [12] to the PCM device and by self-consistently solving also the corresponding circuit equations mixed mode [66, 75]. The used load resistor, in series to the voltage generator, was 4 KOhm while a capacitor of 25 pF was added to the electrical net between the cell top electrode and ground, describing the role of the probe tip capacitance. The capacitance causes a delay in the switching event of about 70 ns. The voltage drop on GST during the pulse plateau is instead determined by the partition between the load resistance and the on-state device resistance of about 1 kΩ.

4.4.2 Thermal Model

To achieve a full description of the PCM programming operation, the heat conduction equation has been self-consistently coupled with the electrical model to describe the thermal heating through the Joule effect. The employed thermal conductivities and specific heats are reported in Table 4.3 [13, 21]. To take into account for the heat transport related to carriers generated in GST at high current densities and temperatures, the Wiedemann-Franz equation has been locally implemented for the chalcogenide, leading to an increase in the thermal conductivity (at about 0.012 W/cm K) in the region with the higher carrier concentration, i.e., at the chalcogenide/heater interface. Moreover, specific heat and thermal conductivity for a typical heater material (TiN) have been taken from literature [56].

Thermal maps can thus be obtained by performing transient simulations on an initially crystalline device for several current values, as reported in Fig. 4.6. In the first frame, the melting condition can be recognized, corresponding to the existence of a thin melted layer of GST close to the interface with the metallic pillar. The transition from the melt to the amorphous phase is actually determined by the

Table 4.3 Thermal conductivities and specific heats for materials in the active region of the device

	cGST	α-GST	Heater	SiO_2
$k_{th}\left(\frac{W}{cm\ K}\right)$	0.005	0.002	0.05–0.15	0.007–0.01
$c_s\left(\frac{J}{cm^3\ K}\right)$	1.25	1.25	1.7	3.1

Adapted with permission from [13]; copyright AIP Publishing LLC 2008

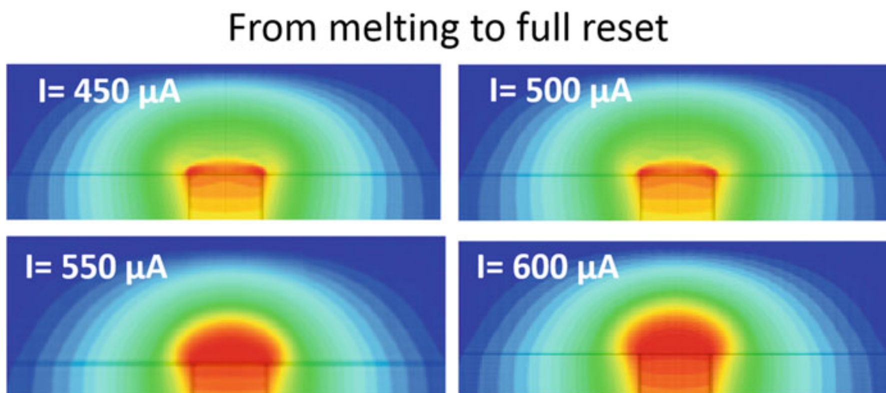


Fig. 4.6 Thermal maps obtained at the end of a 50 ns pulse for the increasing current pulses reported on the frames. *Red color* corresponds to temperatures higher than melting temperature ($T > 888$ K), while the *blue color* represents the room temperature (300 K) (Adapted with permission from [12]; copyright IEEE 2005)

quenching time. In fact, during the falling edge, the GST cools down passing through the crystallization temperature region, thus allowing nucleation and growth to take place with their own kinetics. If the falling edge is very short (10–15 ns), crystallization is practically avoided, so that the hot region becomes *a*GST, resulting in a low conductive state. Further increases in the programming current cause the thermal spot to extend, and the amorphous zone is correspondingly enlarged, still maintaining a hemispherical shape.

4.4.3 Phase Change Algorithm

To account for the phase transitions responsible for the memory effect, a local phase change algorithm has been coupled to the electrothermal equations. The crystallization takes place by the action of nucleation and growth as previously discussed. The first phenomenon refers to the appearance of nuclei of the new phase in a homogeneous matrix of the old phase, while the second refers to the growth of a pre-existing interface between the new and the old phases. The model implemented is a steady-state homogeneous nucleation and growth model that neglects both transient and heterogeneous effects. Due to the complexity of the model that includes self-consistently the electrothermal problem and since we have no experimental evidences that the modeling of those mechanisms is required for PCM device description, we preferred to implement a simplified crystallization model in which possible effects of the transient or heterogeneous nucleation are accounted for by effective crystallization parameters. So, starting from the nucleation rate and growth velocity originally proposed by Peng et al. [21], a local Monte Carlo (MC) algorithm has been implemented and calibrated on experimental data. The grid element dimension was chosen to be comparable with the critical cluster size at the temperature for which nucleation probability shows its maximum. The nucleation and growth probabilities have been calculated starting from the nucleation rate and growth velocity as reported in the *Memory Switching* section. Details on the normalization of nucleation rate and growth velocity as a function of the grid size can be found in [13].

Such probability has been applied for both amorphous-to-crystal and melt-to-crystal transitions. The transition probabilities referred to $\Delta t = 1$ ns are reported in Fig. 4.7. The surface energy for the critical cluster in Eq. (4.11) is taken equal to $\Upsilon = 0.1$ J/m^{-2}. The temperature dependence of bulk free energy has been taken from the literature [29, 63] as extensively explained in [13]. The main parameters related to the phase transition have been reported in Table 4.4 where ΔH_1 is the latent heat of solid-to-liquid transition; ΔH_2 is exothermic heat of amorphous-to-crystalline transformation; E_{an} *e* E_{ag} the activation energy of nucleation and growth, respectively; and T_m and T_g the melting and glass transition temperature, respectively. In the low-temperature range, due to the increase in the atomic mobility with temperature inside the melted GST, both nucleation and growth are enhanced with temperature. On the contrary, for temperatures close to the melting

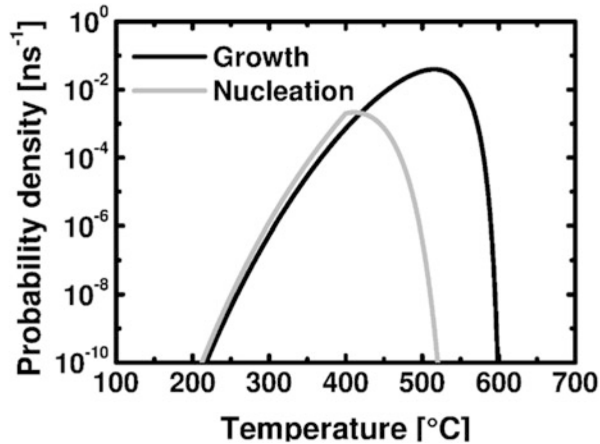


Fig. 4.7 Probability densities for unit time for nucleation and growth mechanism implemented in the Monte Carlo model (Refs. [5, 11, 21]). Note that both nucleation and growth are characterized by a bell-shaped temperature dependence (Reprinted with permission from [5]; copyright IEEE 2004)

Table 4.4 Key numerical values to describe the phase change model for amorphous and crystal transitions

Properties	Value	Properties	Value
E_{an} (eV)	2.6	E_{ag} (eV)	2.6
$\Delta H_1 (J\ \mathrm{cm}^{-3})$	418.9	$\Delta H_2 (\mathrm{J\ cm}^{-3})$	218.5
T_m (K)	888	T_g (K)	673

Adapted with permission from [13]; copyright AIP Publishing LLC 2008

one, both nucleation and growth are turned off with the functional relationship previously discussed in the Sect. 4.3.

4.4.4 Results

The numerical implementation of the full self-consistent electrothermal phase change model relies on the adoption of two different loops. At each time step, an inner loop loads the current phase distribution and self-consistently solves the electrothermal problem, evaluating the temperature at each grid node. Each simulation covers the typical electronic time scale, thus ranging from about 10 ps to about 1 ns. In the outer loop, the nucleation and growth probabilities are thus calculated from the temperature map and modified from nanoseconds to much longer times according with the temperature variations, trying to minimize the computational cost. At the end of the transient simulation, a final phase distribution is achieved, and its resistance can be evaluated.

With the proposed computer model, several full-coupled programming and reading operations can be simulated. The calculations were performed by assuming the numerical values reported in Tables 4.1, 4.2, 4.3, and 4.4.

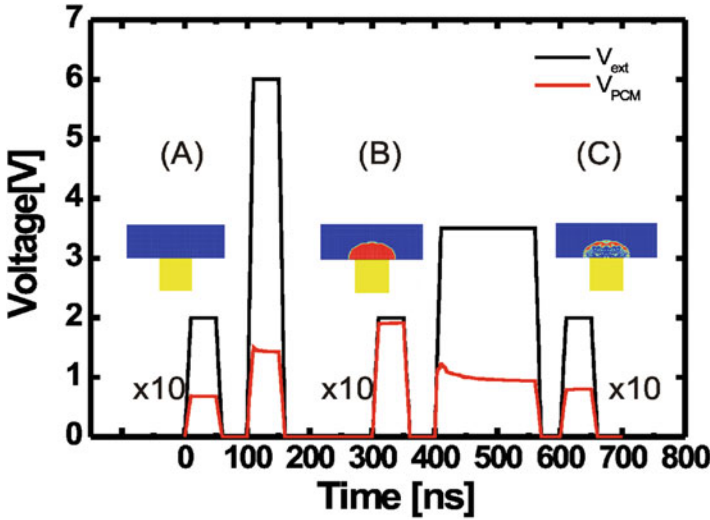


Fig. 4.8 Transient simulation performed by applying a sequence of reading and programming voltage pulses. The first voltage pulse allows to read the initial set state (**a**) of the bit (for clarity reading pulses were multiplied by 10 to make them visible on the same scale of programming pulses). The second pulse carries the bit in the reset state (**b**) as highlighted by the subsequent reading pulse. Finally, a set pulse is applied, and the crystalline state in the inset (**c**) is correctly detected (Reprinted with permission from [13]; copyright AIP Publishing LLC 2008)

Figure 4.8 shows the effect of several pulses on a bit in the initial crystal phase. Since the load resistance has been chosen comparable to the set one, the voltage across the device is practically a half of the external voltage when the bit is in the set state. The first pulse thus correctly detects the status of the bit. The second voltage pulse is high enough to carry the bit to the reset state. In fact, the subsequent reading pulse detects a voltage across the device equal to the external one, demonstrating that the device has been carried to a high resistance level. The fourth pulse then is a programming pulse with an appropriate voltage level, which is able to partially crystallize the amorphous spot as underlined by the final reading pulse. With this approach, it has been possible to investigate and understand the phase distribution of programmed levels. It is worth noting that the electrical model qualitatively reproduces the current filament formation effects observed during threshold switching [27] and that the final phase distribution results from the full coupling among the electrothermal and phase change models, thus promoting phase transitions in the hottest regions, i.e., in the current filaments as depicted in Fig. 4.9. The amorphous-to-crystal transition is thus a localized process following some preferential paths, depending on current density and time, while the nature of crystal-to-amorphous transition is practically isotropic, thus proceeding with hemispherical shape [5]. By properly choosing the sequence of programming and reading pulses, the so-called programming characteristic can be obtained. In Fig. 4.10, low-field resistances are plotted as a function of the programming current. The curve is

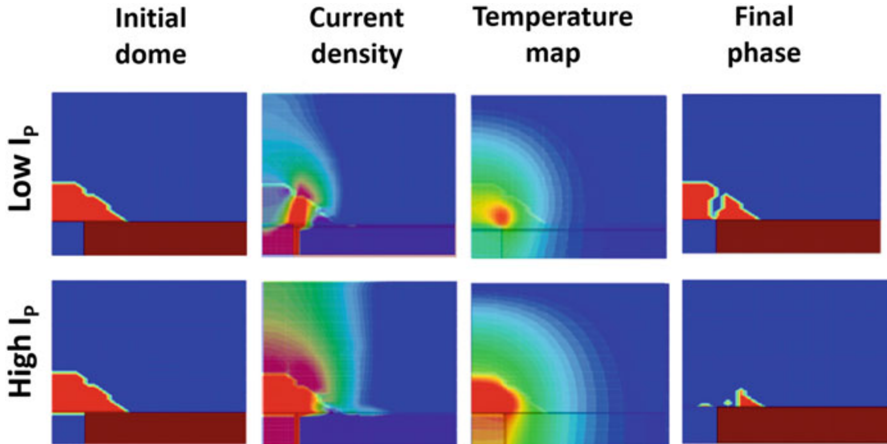


Fig. 4.9 In the first row, starting from an initial phase distribution reported on the first frame on the left, a low programming current pulse has been applied to the device. The electronic switching leads to a localized current path that spikes in the highest field region. A localized self-heating is promoted only in the high current region causing a localized crystal path. In the second row, starting from the same initial phase distribution reported on the first frame on the left, when a higher current pulse is applied to the device, the high current density region is much larger than in the previous case, leading to an increased size of the crystallized zone (Adapted with permission from [5]; copyright IEEE 2004)

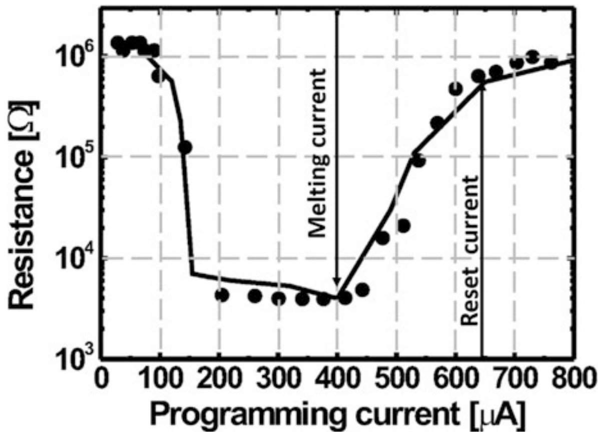


Fig. 4.10 Programmed resistance obtained by applying to an amorphous bit programming pulses of 150 ns and increasing current amplitude. For programming current lower than 150 μA, the amorphous bit does not change its state. For programming current ranging from 150 to 400 μA, the initial amorphous bit is carried in the crystalline state. Finally, for programming currents higher than 400 μA, the initial amorphous bit starts to melt, and for programming current of about 600 μA, the amorphous resistance is totally recovered. Also, in this case, a good agreement between experimental data (*hollow circles*) and calculations (*solid line*) is achieved (Reprinted with permission from [5]; copyright IEEE 2004)

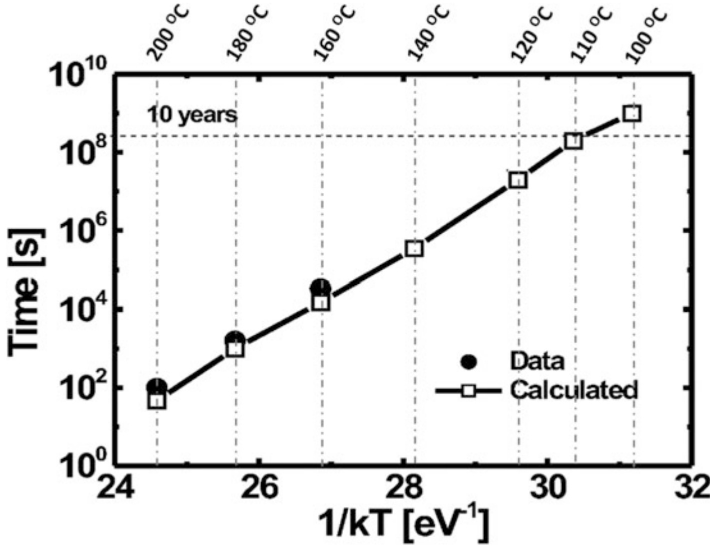


Fig. 4.11 Experimental and simulated crystallization times (defined as the time required for the initial amorphous bit to reach a resistance of ten times the set one), during constant temperature annealing experiments, as a function of reciprocal temperature. Note that an annealing temperature lower than 110 °C ensures a crystallization time longer than 10 years (Reprinted with permission from [5]; copyright IEEE 2004)

obtained starting from an amorphous bit and applying increasing programming pulses with a 150 ns width and a 10 ns falling edge. For $I < 100$ µA, the cell resistance remains close to the amorphous one.

Higher current densities promote the GST crystallization process, thus causing resistance values to drop down [5]. Further increases in the programming current results in the melting condition (400 µA), while at about 600 µA, the amorphous volume saturates. The same figure reports the simulated *R-I* curve, resulting in good agreement with experimental data. To check the model, the times requested for an initially amorphous bit to crystallize during constant temperature annealing, i.e., failure times, have been also calculated for several temperatures and compared with experimental data in the Arrhenius plot depicted in Fig. 4.11. Each failure time is defined as the time for which the resistance reaches a reference value, i.e., ten times the minimum set resistance. Also in this case, good agreement with experiments has been achieved, and both experiment and simulation show that working temperatures lower than 110 °C are requested to guarantee a time of crystallization of 10 years. However, unwanted heating of the device may result in disturbing bits in the amorphous state [9]. The model could also be easily employed to make 3D cross-talk simulations with the purpose of evaluating the temperature induced by a programming operation on the adjacent amorphous bits [6, 77].

4.5 Conclusions

The present work reports a review of the main physical ingredients required for PCM device simulations. Starting from the electronic equilibrium and transport properties modeling of both crystal and amorphous phases, particular emphasis has been devoted to the TS and MS mechanisms. These effects, in fact, have been demonstrated to have a major impact on the PCM device functionality, and they cannot be neglected in PCM modeling. The proposed TS model relies on a bulk generation-recombination balance related to the huge density of defects in the amorphous chalcogenide. It has the advantage to be able to describe the main observed experimental evidences, being at the same time compatible with standard simulation tools for semiconductor. The developed numerical tool includes electrothermal and phase transition physics. The model implements a steady-state local nucleation and growth algorithm able to take into account percolative effects and neglects the effect of surfaces and boundary materials on the crystallization kinetics (heterogeneous effects). Furthermore, the set of equations implemented in the computer model do not have the capability to describe the latent heat during the phase transitions. Since until now there are no experimental evidences that the transient effects, heterogeneous effects, and latent heat descriptions are required for the PCM behavior description, we preferred to implement a simpler crystallization model. The tailoring of the electrothermal and phase transition model parameters have been carried out by comparison with the experimental data acquired on different device architectures and technology nodes. A reliable set of parameters has been then defined leading to a comprehensive numerical tool suitable for PCM cell design and to support the development of this technology for nonvolatile memory applications.

Bibliography

1. A.C. Faleiros et al., Mater. Res. **3**(3), 51 (2000)
2. A.E. Owen, J.M. Robertson, IEEE Trans. Electron Devices **20**(2), 105 (1973)
3. A.E. Owen et al., J. Non-Cryst. Solids **32**(1-3), 29 (1979)
4. A. Kolobov et al., Appl. Phys. Lett. **91**(2), 021911 (2007)
5. A.L. Lacaita et al., IEEE IEDM Tech. Dig. 911 (2004)
6. A. Pirovano et al., IEEE IEDM Tech. Dig. 699 (2003)
7. A. Pirovano et al., IEEE Trans. Electron Devices **51**(3), 452–459 (2004a)
8. A. Pirovano et al., IEEE Trans. Electron Devices **51**(5), 714–719 (2004b)
9. A. Pirovano et al., IEEE Trans. Device Mater. Reliab **4**(3), 422–427 (2004c)
10. A. Redaelli et al., IEEE Electron Device Lett **25**(10), 684–686 (2004)
11. A. Redaelli et al., IEEE IEDM Tech. Dig. 761 (2005a)
12. A. Redaelli et al., SISPAD Proc. 279–281 (2005b)
13. A. Redaelli et al., J. Appl. Phys. **103**, 111101 (2008)
14. A. Redaelli et al., IEEE Trans. Electron Devices **53**(12), 3040–3046 (2006)
15. E.J. Yoffa, D. Adler, Phys. Rev. B **15**, 2311 (1977)
16. M. Avrami, J. Chem. Phys. **7**, 1103 (1939)

17. M. Avrami, J. Chem. Phys. **8**, 212 (1940)
18. M. Avrami, J. Chem. Phys. **9**, 177 (1941)
19. B.S. Lee et al., J. Appl. Phys. **97**, 093509 (2005)
20. C.N. Afonso et al., Appl. Phys. Lett. **60**, 3123 (1992)
21. C. Peng et al., J. Appl. Phys. **82**, 4183 (1997)
22. S. Caravati et al., Appl. Phys. Lett. **91**, 171906 (2007)
23. C. Jacoboni et al., Solid State Electron. **84**, 90–95 (2013)
24. J.W. Christian, *The Theory of Transformation in Metals and Alloys* (Oxford, Pergamon, 1965)
25. Y. Okuto, C.R. Crowell, Solid State Electron. **18**, 161–168 (1975)
26. D. Adler et al., Rev. Mod. Phys. **50**, 209 (1978)
27. D. Adler et al., J. Appl. Phys. **51**, 3289 (1980)
28. D. Ielmini, Y. Zhang, IEEE IEDM Tech. Dig. 401 (2006)
29. D. Turnbull, M. Cohen, J. Chem. Phys. **34**, 120 (1961)
30. E. Piccinini et al., J. Appl. Phys. **112**, 083722 (2012)
31. E. Piccinini et al., Appl. Phys. Lett. **104**, 022101 (2014)
32. D. Turnbull, J.C. Fisher, J. Chem. Phys. **17**, 71 (1949)
33. G.C. Sosso et al., Phys. Status Solidi B **249**(10), 1880–1885 (2012)
34. G. Wicker, *A Comprehensive Model of Submicron Chalcogenide Switching Devices* (Wayne State University, Ph.D Thesis, 1996)
35. G.C. Sosso et al., J. Phys. Chem. B **118**(47), 13621–13628 (2014)
36. H. Fritzsche, S.R. Ovshinsky, J. Non-Cryst. Solids **2**, 393 (1970)
37. H. Iwasaki et al., Jpn. J. Appl. Phys. **32**(1), 5241 (1993)
38. I. Friedrich et al., J. Appl. Phys. **87**, 4130 (2000)
39. D. Ielmini, Phys. Rev. B **78**, 035308 (2008)
40. J.C. Fisher et al., J. Appl. Phys. **19**, 775 (1948)
41. J. Coombs et al., J. Appl. Phys. **78**, 4906 (1995a)
42. J. Coombs et al., J. Appl. Phys. **78**, 4918 (1995b)
43. J. Kalb et al., J. Appl. Phys. **93**, 2389 (2003)
44. J. Park et al., Jpn. J. Appl. Phys. Part 1 **38**, 4775 (1999)
45. J. Akola, R.O. Jones, Phys. Rev. B **76**, 235201 (2007)
46. K. Yusu et al., Proc. EPCOS (2005)
47. M. Kastner, Phys. Rev. Lett. **28**, 355 (1972)
48. S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1976)
49. N. Kolmogorov, Biol. Bull. Acad. Sci. USSR **3**, 3551 (1937)
50. A. Kolobov et al., Nat. Mater. **3**, 703–708 (2004)
51. E.R. Meinders, M.H.R. Lankhorst, Jpn. J. Appl. Phys. **42**(Part 1, Number 2B), 809 (2003)
52. M. Chen et al., Appl. Phys. Lett. **49**, 502 (1986)
53. M.H. Cohen et al., Phys. Rev. Lett. **22**, 1065 (1969)
54. M.H.R. Lankhorst et al., Jpn. J. Appl. Phys. **42**(Part 1, Number 2B), 863 (2003)
55. M. Simon et al., J. Appl. Phys. **108**, 064514 (2010)
56. M.V. Barsoum et al., J. Appl. Phys. **87**, 8407 (2000)
57. M. Wuttig, N. Yamada, Nat. Mater. **6**, 824–832 (2007)
58. M. Rudan et al., IEEE Trans. Electron Devices **58**(12), 4361–4369 (2011)
59. N. Yamada, T. Matsunaga, J. Appl. Phys. **88**, 7020 (2000)
60. W.A. Johnson, R.F. Mehl, Trans. AIME **135**, 416 (1939)
61. N. Yamada et al., J. Appl. Phys. **69**, 2849 (1991)
62. S.R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968)
63. A.E. Owen, *Amorphous Solid and the Liquid State* (Plenum, New York, 1985)
64. P. Fantini et al., Appl. Phys. Lett. **88**, 263506 (2006)
65. P. Fantini et al., *Micron TLP Journal* (Micron Internal TLP Journal, Boise, 2015)
66. R.M. Kielkowski, *Inside Spice* (McGraw-Hill, New York, 1998)
67. W. Van Roosbroeck, Phys. Rev. Lett. **28**, 1120 (1972)
68. S. Caravati et al., J. Phys. Condens. Matter **21**, 255501 (2009)

69. S.A. Baily et al., Solid State Commun. **139**, 161 (2006)
70. S. Gabardi et al., Phys. Rev. B **92**, 054201 (2015)
71. S. Lai, T. Lowrey, IEEE IEDM Tech. Dig. 803 (2001)
72. S.M. Wiggins, al. Appl. Phys. Lett. **84**, 4445 (2004)
73. S. Tyson et al., Aerospace Conference Proceedings (unpublished) **5**, 385 (2000)
74. *Sentaurus Device User Guide, version Z, Synopsis*, (2007)
75. T. Quarles et al., *SPICE 3 Version 3F5 User's Manual*, 2nd edn. (Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, 1994)
76. U. Russo et al., IEEE Trans. Electron Devices **53**(12), 3032 (2006)
77. U. Russo et al., IEEE Trans. Electron Devices **55**(2), 506 (2008)
78. V.G. Karpov et al., Appl. Phys. Lett. **92**, 173501 (2008)
79. V. Weidenhof et al., J. Appl. Phys. **89**, 3168 (2001)
80. W. Welnic et al., Nat. Mater. **5**, 56–62 (2006)
81. A.C. Warren, IEEE Trans. Electron Devices **20**, 123 (1973)
82. D. Wright, S. Senkader, J. Appl. Phys. **95**, 504 (2004)
83. C.A. Volkert, M. Wuttig, J. Appl. Phys. **86**, 1808 (1999)
84. Z. Sun et al., Phys. Rev. Lett. **96**, 055507 (2006)
85. D. Ielmini, Y. Zhang, J. Appl. Phys. **102**, 054517 (2007)

Chapter 5
PCM Main Reliability Features

Robert Gleixner

5.1 Introduction

To effectively realize the performance benefits of PCM over existing technologies, the memory cells must be able to operate reliably over the field life of the product. When compared to DRAM and NAND flash memory, PCM's capabilities drive reliability requirements that in many cases fall between these two established technologies. The much lower read and write latency, combined with random access to the array (not block-based as in flash memory), drive a requirement for higher write endurance capability and resistance to read disturbs than NAND flash. While the endurance requirements could be reduced (as is done in NAND flash) with wear leveling and read caching, such an implementation would reduce the effective performance of the PCM device and is therefore undesirable. At the same time, the PCM write latency is significantly higher than DRAM, and therefore the write endurance requirement is far lower than DRAM. Table 5.1 below lists the effective latencies of NAND flash, DRAM, and PCM and how this translates into the required write endurance.

While the operational requirements (reads and writes) fall between NAND and DRAM, the fact that PCM is targeting nonvolatile memory applications imposes data retention requirements that are similar to NAND flash. For typical consumer electronics applications, this requires that the component supports an operating temperature of 85 °C for "reasonable" periods of time (days to weeks) and much longer times (5–10 years) at ~50 °C. These specifications may be insufficient for some applications, for example, in automotive, military, and space applications, which require either extended retention time, higher temperature capability, or

R. Gleixner (✉)
Micron Technology, Inc., Boise, ID, USA
e-mail: rjgleixn@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_5

Table 5.1 Comparison of read and write performance between NAND flash memory, DRAM, and PCM

Target specifications	NAND flash	DRAM	PCM
Read latency	10 μs (page)	~10 ns	~100 ns
Write latency	1 ms (page)	~10 ns	100 ns–1 μs
Endurance	10^3–10^5	$>10^{15}$	10^6–10^8

both. Moreover, in military and space applications, the devices typically must retain data in the presence of high-radiation environments.

With these requirements in mind, the purpose of this chapter is to review the ability of PCM to meet the requirements of a reliable nonvolatile memory product. The chapter will focus on the intrinsic behaviors of the PCM material that limit its ability to either change state when required (which corresponds to write endurance) or unintentionally change state (the case for disturb and retention loss).

5.2 Overview of Failure Modes and Mechanisms

PCM reliability failures are primarily governed by two mechanisms. The first is the inherent instability of the amorphous phase (which thermodynamically prefers to be crystalline), and the second is a change in the active material composition over time. While less common, there are special cases where the cell can fail due to unintentional amorphization. This could result from an electrical overstress, which should be prohibited by the design of the array circuitry (and will not be discussed here), or exposure to high doses of energetic ions (which will be briefly discussed at the end of the chapter).

Within the primary mechanisms, the failures of PCM are typically classified by the usage specifications that are impacted. These include data retention, cycling endurance, read disturb, and thermal disturb (often called proximity or write disturb). Figure 5.1 contains a cartoon that depicts these failures in what is commonly referred to as a “lance” or a “mushroom” type of PCM cell. For further details of the device cell architecture, see Chap. 9. In this case the cell is formed by contacting a heating electrode (gray) with the chalcogenide (pink or red color). In the chalcogenide region, the crystalline phase is pink colored, and the amorphous phase is red. Since the crystalline state is stable, the bulk of the chalcogenide in this cell architecture will take this state. When the device is programmed to RESET, only the small volume directly above the heater is converted to the amorphous phase. Failure of the cell is depicted as a change in this region away from sufficiently large, uniform condition. While other PCM cell architectures may vary in the details – it is of course possible to have heating from several sides, or to confine the chalcogenide such that the entire volume is amorphized – they typically share these same mechanisms.

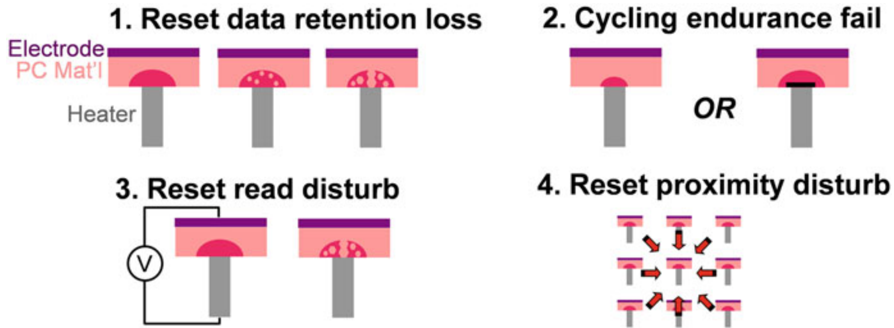


Fig. 5.1 Cartoon depicting the reliability failure modes typically observed in PCM

Considering the failure modes shown in Fig. 5.1, three of the four share the same mechanism and primarily impact the RESET state. RESET data retention loss occurs when the amorphous region crystallizes during nonoperation of the cell. As the crystallization process is thermally activated, this failure mode is highly sensitive to the ambient temperature of the device. The RESET read disturb failure mode shares the same mechanism but is separated from data retention by the inclusion of an applied bias. (The applied bias in this case could just be considered an accelerating factor for the crystallization process, which will be discussed later in this chapter.) And finally, proximity disturb is also a disturb of the RESET state but in this case is due to a very localized increase in temperature due to writing operations performed on the neighboring cells.

While the abovementioned failures share the same mechanism, cycling endurance fails have a broader variation in the underlying mechanisms. In addition, the capability of the cell to withstand write cycles without degrading is much more dependent on the cell architecture, electrode materials, and specific manufacturing flow. The cartoon in Fig. 5.1 shows two typically reported mechanisms for cycling endurance fails. In the first case, the failure occurs due to the amorphous region not being sufficiently large to increase the cell resistance to the level required for differentiating it from the SET state. This could occur due to degradation of the heating element such that the applied current is no longer sufficient to melt the required volume during the RESET operation. Alternatively, it could also result from a change in the chalcogenide itself that increases the melting temperature (so again, the RESET operation cannot achieve the required resistance). The second cartoon depicts the case where the cell would be stuck in the high-resistance state, which could result from delamination or a void forming at the electrode interface.

5.3 Data Retention

In a PCM device, data retention failures occur due to a shift in the resistance of the cell with time. Since the crystalline phase is stable, the primary risk is driven by the relative instability of the amorphous RESET state. There are two mechanisms by which RESET resistance changes with time, as shown in Fig. 5.2. After the cell is programmed, the resistance first increases due to the mechanism referred to as resistance drift. As the crystalline SET state does not drift, this has the effect of increasing the resistance window between the states and is generally not a concern. (While this is true for single-level cell operation, drift is a significant concern in multilevel cell operation. Methods to mitigate this issue will be addressed later in the chapter.) As time progresses, the amorphous state begins to crystallize and the resistance declines. Eventually the resistance will fall below the level that demarcates a RESET vs. a SET cell, and data loss will occur.

Data retention is perhaps the most studied reliability mechanism in PCM, as it is based on an intrinsic compromise in the material. To achieve low write latency, the material must be able to crystallize within hundreds of nanoseconds during the programming operation. In the absence of programming, crystallization must be suppressed for years. The popularity of the GeSbTe system for PCM is largely due to its ability to meet these requirements. As discussed in Chap. 4, the crystallization mechanism is a thermally activated process. A high thermal activation energy of crystallization is a required property of the material, allowing for programming speeds in the tens of nanoseconds (here considering RESET to SET) by heating the cell to several hundred degrees Celsius but at the same time enabling many years of data retention at typical product operating temperatures.

To estimate the data retention capability of PCM, RESET cell resistance is measured as a function of time at multiple temperatures. Figure 5.3 shows distributions of cell resistance as a function of bake stress time, from Gleixner et al.

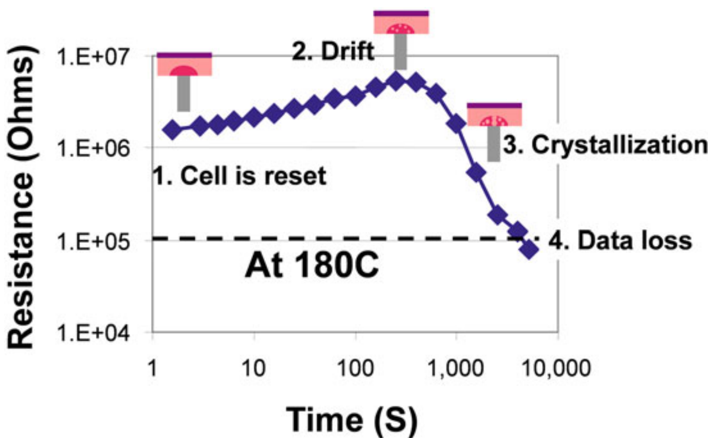


Fig. 5.2 Representative PCM resistance vs. time curve, in this case at elevated temperature

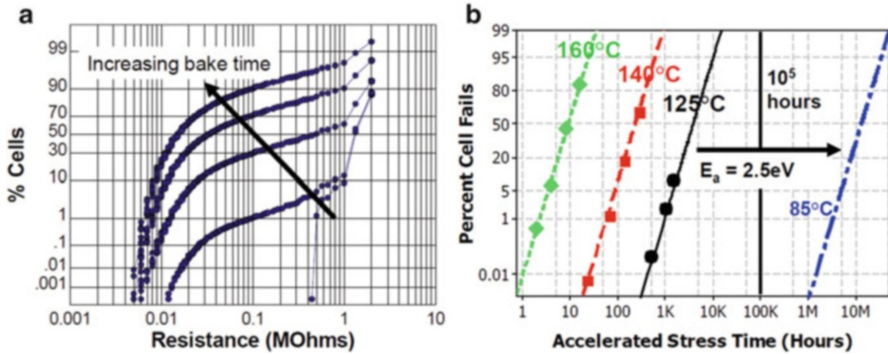


Fig. 5.3 (**a**) shows the cell resistance distributions as a function of accelerated bake time for PCM cells. From these distributions, the fraction of cells failing the RESET resistance requirement (100 kΩ) is plotted as a function of bake time and at multiple temperatures. This is then extrapolated to an 85 °C product operating temperature (Both plots are from Gleixner et al. [1])

[1]. These data show the significant variation in crystallization time across an array of cells, to the extent that a significant portion of the cells have reached their SET resistance (less than 10 kΩ), while the majority of the cells are still showing no sign of resistance loss (greater than 1 MΩ). This statistical variation is an important aspect of PCM that must comprehended, as product reliability will be governed by the fastest crystallizing cells. To estimate this variation, data such as those shown in Fig. 5.3a are collected at multiple temperatures. The fraction of cells failing the RESET resistance requirement is then extracted, and the failure rates vs. time and temperature are plotted as in Fig. 5.3b. In this case, a lognormal distribution is used to fit the failure rate vs. time curves and appears to represent the failure rate distribution very well. Other analyses have proposed that a Weibull distribution or a log-logistic distribution could better represent the data (a common debate in the reliability community that most likely depends on the detail of the dominant source of variation). The Arrhenius model is used to scale these data from the elevated testing temperature to typical product use temperatures, which is given by

$$t \propto e^{\frac{E_A}{k_B T}}$$

Here t is the time to crystallize (RESET failure), T is temperature in Kelvins, E_A is the thermal activation energy, and k_B is Boltzmann's constant. Figure 5.3b shows both the fit to this model with a single E_A as well as the resulting prediction of the cell failure rate vs. time at 85 °C. As the data indicate, the high activation energy (2.5 eV) lead to a retention capability that is well over 10^5 hours at 85 °C.

While the prior analysis indicates that the intrinsic retention can support commercial products, the broad distribution of failure times is an aspect of PCM that must be understood. Perhaps the most important component of this variation to resolve is whether it is due to random variations in the crystallization process itself or differences in the cell-to-cell structure. The presumption is that if the earliest

Fig. 5.4 Variation in response of individual PCM cells to repeated RESET and bake operations (From Gleixner et al. [1])

failing cells are structurally different, then retention could be managed by manufacturing screens (e.g., replacing these cells with redundant elements). On the other hand, if the variability is due mainly to a statistical variation in crystallization time, then replacing cells is not a robust solution (since the earliest failing cells will not be consistent). Figure 5.4. shows the results of an experiment where cells were programmed (RESET) and baked four consecutive times. Individual PCM cells were chosen with the common property in that they had failed Bake 2, and then their behavior was tracked across the other bakes. These data show a low level of consistency bake-to-bake, with none of the cells failing any of the other bakes and a wide variation in their post-bake resistance. This suggests that the crystallization mechanism is stochastic in nature and at least partly responsible for the variation.

To further understand this variation requires a model of the crystallization mechanism. As was discussed in Chap. 4, the crystallization process in chalcogenides is comprised of both a nucleation phase and a growth phase. While the time spent to nucleate could be greatly dependent on a number of factors (very fast if “seed” crystals are available and much slower if homogenous nucleation is required), in typical PCM cells with GeSbTe alloys, the nucleation and growth rates are expected to be comparable at temperatures less than 250 °C [2]. Figure 5.5 shows the results of simulating crystallization in GeSbTe (from Redaelli et al. [3]). Figure 5.5a shows data similar to Fig. 5.4, where a single cell demonstrates a variability in crystallization rate with repeated bake tests. Figure 5.5b shows the results of a model where crystallites nucleate within the bulk of the amorphous region and grow and where the calculated resistance loss responds to the extent of transformation. Using the measured data to calibrate the model, this analysis showed the ability to simulate the observed variation in crystallization time. Furthermore, it suggests that at lower temperatures (<100 °C), the crystallization time is dominated by the growth process as the critical nucleation radius becomes

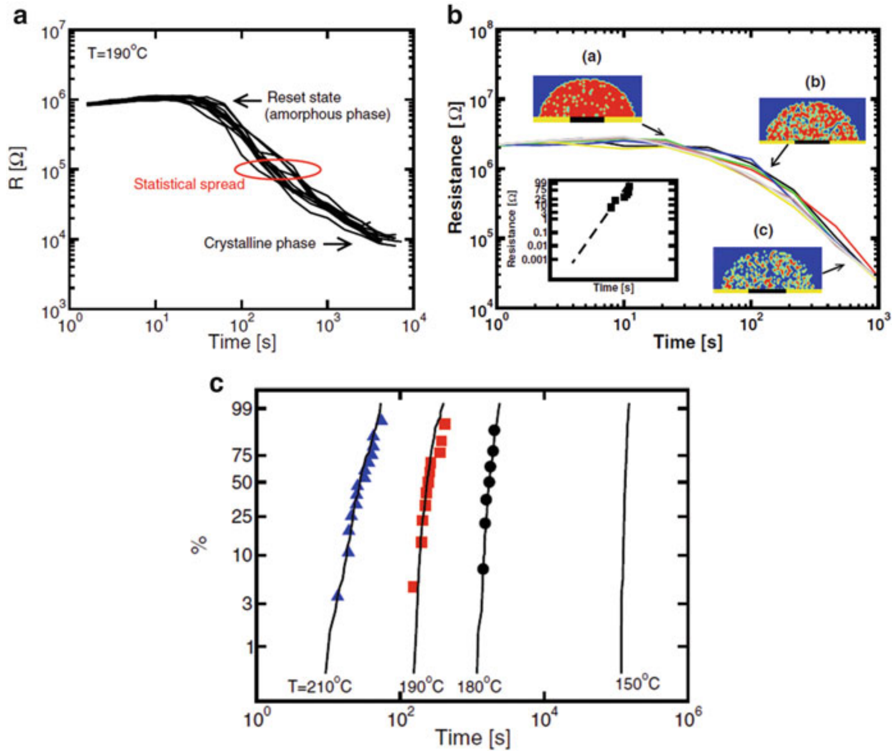


Fig. 5.5 (**a**) shows the measured variation in PCM crystallization on a single cell, (**b**) is a simulation of PCM amorphous phase crystallization and the corresponding evolution of cell resistance, and (**c**) is a comparison between data (the points) and simulation (the fitted lines) as a function of temperature (All graphs are from Redaelli et al. [3])

very small and many nuclei are available. Since the variation in the overall crystallization process is a function of variation in both the nucleation and growth times, as the nucleation time is reduced, the process will be increasingly dominated by just the growth process. This should produce a reduction in variation of crystallization time at lower temperature.

Beyond the within-cell statistical variation, there are number of sources of potential cell-to-cell variation. One of these is compositional fluctuations of the chalcogenide material. While the impact of composition on data retention will be discussed later in this chapter, an analysis of local variation suggests that nanometer-scale fluctuations could produce the observed variation in cell behavior [4].

Since the crystallization rate is determined by the characteristics of the amorphous region, the RESET programming parameters are also expected to impact data retention. Gleixner et al. considered the case of what initially appeared to be extrinsic, early failing cells in bake retention testing [1]. Further analysis on these cells showed that the root cause was an insufficient RESET operation, and

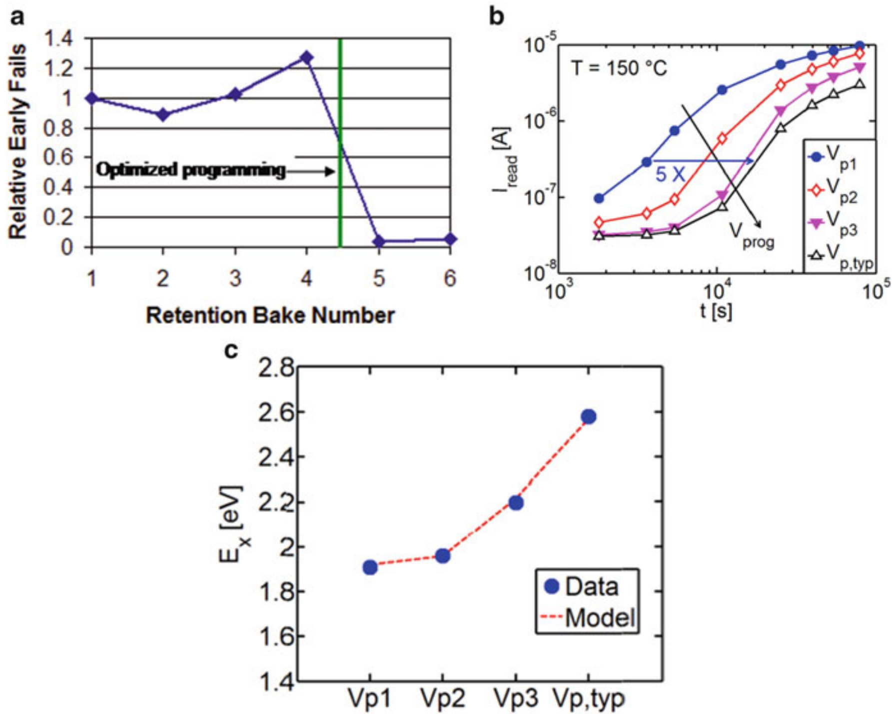


Fig. 5.6 (**a**) shows the impact of optimized RESET programming on reducing an early population of retention fails. The *x*-axis is a series of repeated retention bakes on the same cells [1]. Plots (**b**) and (**c**) show the results of data retention tests on a 45 nm PCM cell with different programming voltages [5]. Here we see that both the crystallization rate can be reduced (**b**) and activation energy increased (**c**) by increasing the programming voltage

optimizing the programming conditions removed this failure mode (Fig. 5.6a). At the intrinsic level, this was examined on a 45 nm PCM cell by Rizzi et al. [5]. Here they varied the programming voltage (which corresponded to the RESET current) and were able to decrease the retention time by a factor of five (Fig. 5.6b). At the lowest programming voltage (V_{p1}), the initial RESET resistance is already degraded, so this may not be a surprising result. But at the higher programming voltages, we see a significant increase in retention with a very small change in the initial resistance. Furthermore, the activation energy of the crystallization process continues to increase with increasing programming voltage (Fig. 5.6c). This suggests that while not detectable by cell resistance alone, the amorphous region characteristics are continuing to evolve with higher levels of programming current.

While the prior studies show data retention to be adequate for typical commercial PCM applications, there is significant interest in applications where higher temperature retention is required. Automotive applications, for example, typically require 125 °C for 10 years. Other applications may require the ability to program the device before a high-temperature soldering operation (to mount the component

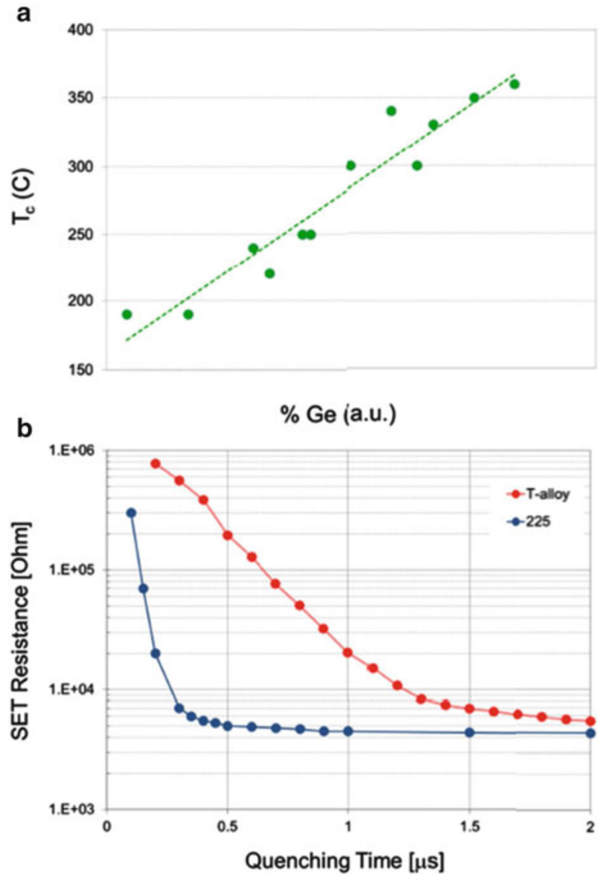


Fig. 5.7 (**a**) shows the increase in crystallization temperature with Ge concentration, while (**b**) shows the decrease in SET speed in a Ge-rich alloy (From Zuliani et al. [6])

to a circuit board), which requires several minutes of time at temperatures greater than 250 °C. As was discussed earlier in this chapter, the GeSbTe system shows crystallization characteristics that make it an ideal material for a nonvolatile memory. But achieving both high SET speed and very-high-temperature retention is difficult (if not impossible) with $Ge_2Sb_2Te_5$. At the same time, due to the high degree of success with integrating GeSbTe into scaled PCM cells, it is highly desirable to work within this system as opposed to moving to an alternative chalcogenide. With this in mind, two methods have typically been investigated for improved retention: altering the stoichiometry or doping the chalcogenide with different materials. Figure 5.7a shows the increase in crystallization temperature of GeSbTe as a function of Ge concentration (from Zuliani et al. [6]). This shows the ability to increase the crystallization temperature by over 150 °C. While this improvement is significant, it results in a reduction of SET write performance. This is shown in Fig. 5.7b, where the time required to SET the Ge-rich alloy is about 10x longer than the $Ge_2Sb_2Te_5$. It does, however, achieve the desired result of increasing data retention. The improvement in high-temperature data retention is

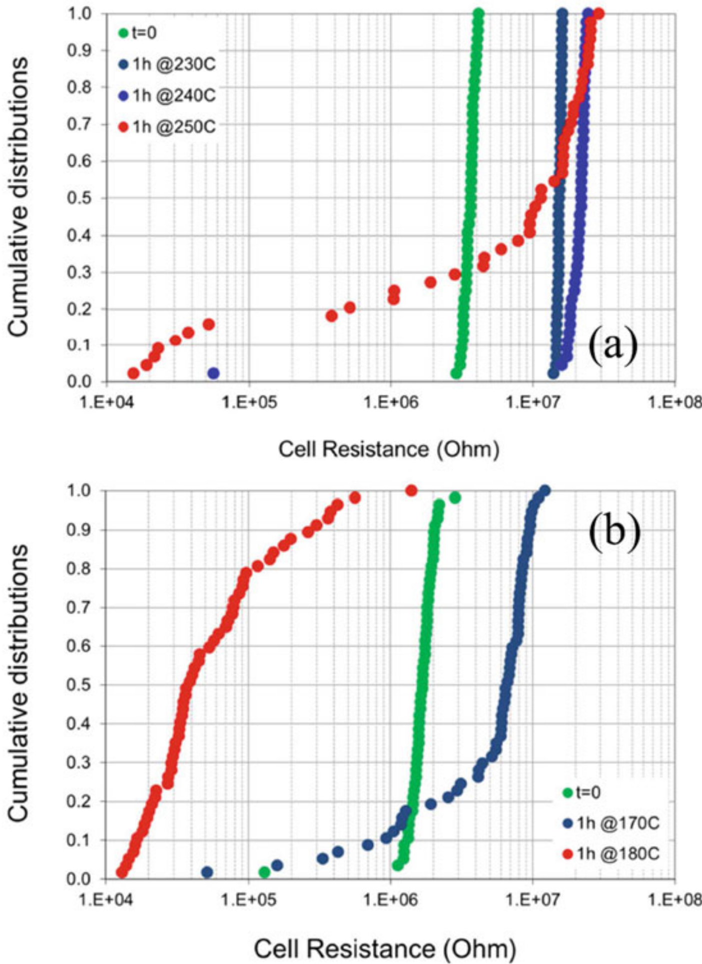


Fig. 5.8 Data retention capability with (**a**) a Ge-rich alloy and (**b**) $Ge_2Sb_2Te_5$ (From Zuliani et al. [6])

shown in Fig. 5.8. In the case of a Ge-rich alloy (Fig. 5.8a), there is little resistance change through 250 °C with a 1 hour bake. In standard $Ge_2Sb_2Te_5$ (Fig. 5.8b) crystallization is apparent after 1 hour at 170 °C.

An alternate approach to varying the chalcogenide composition has been to include dopants. While many different dopants have been tried, two commonly reported are carbon and nitrogen. Figure 5.9 shows the results of a study that included both increasing Ge concentration as well as adding nitrogen and carbon dopants. These show similar results as above, where increasing the Ge concentration increases the crystallization temperature (effectively increasing the retention time at operating temperature). Adding N_2 or C as dopants further increases this

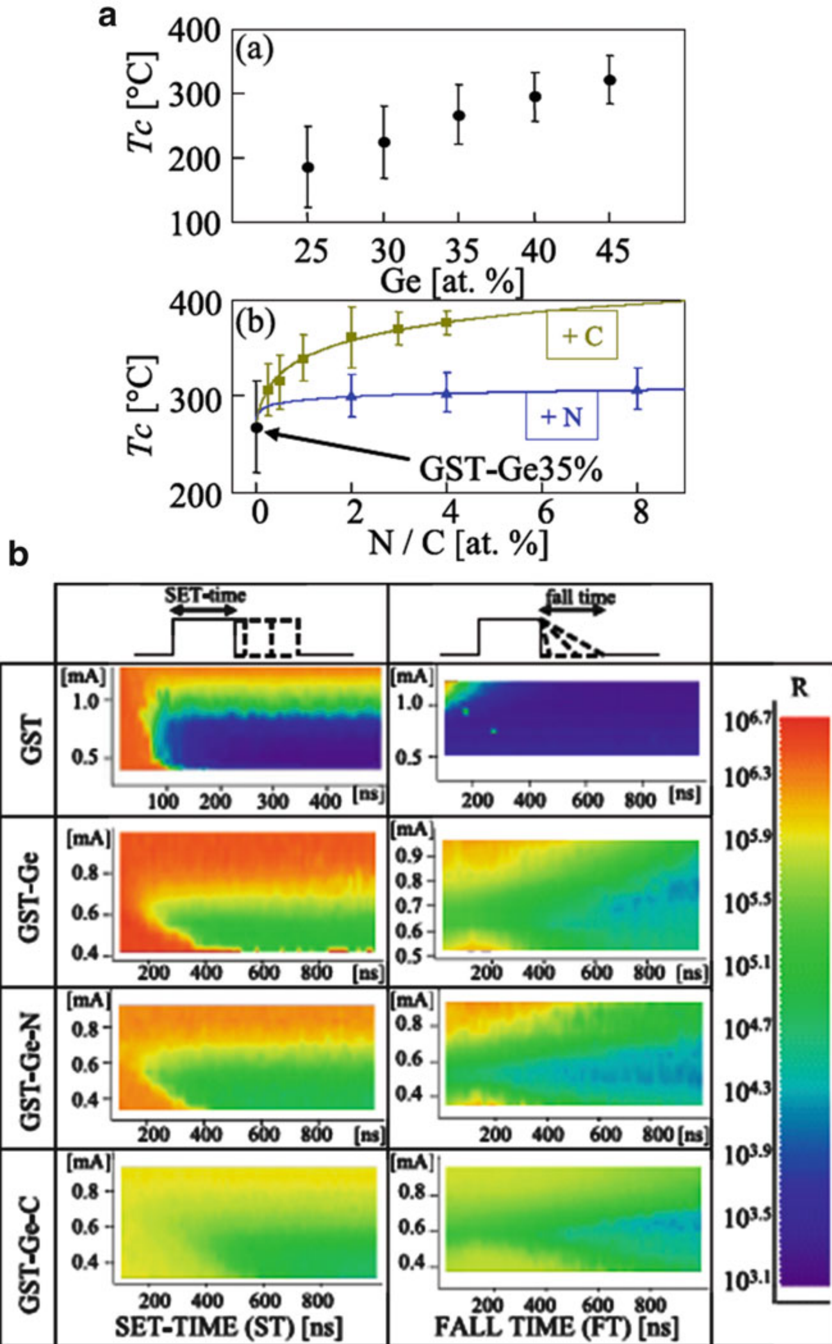


Fig. 5.9 (**a**) shows an increase in crystallization temperature with Ge concentration and with the nitrogen and carbon doping. (**b**) shows the corresponding increase in SET resistance and decrease in SET speed as compared to GST (From Navarro et al. [7])

temperature, as is shown in Fig. 5.9a. While these alloys all show degraded SET characteristics as compared to $Ge_2Sb_2Te_5$ (Fig. 5.9b), they could provide an effective solution in applications where high-temperature retention is required but SET speed is not a critical parameter.

Due to the high activation energy for PCM crystallization, product qualification conditions for PCM components should be carefully considered to ensure that the testing is capturing the field risk of failure. Most commercial memory devices today are expected to pass the JEDEC JESD-47 standard. Prior to 2012, the data retention stress requirements in this standard were based on the charge loss mechanism in floating gate (flash memory) devices, which has a thermal activation energy of 1.1 eV. The uncycled high-temperature data retention test was specified as a 125 °C bake for 1000 hours, based on a goal of continuous 70 °C operation for 10 years. With the higher activation energy for PCM (2.5 eV), this bake condition would correspond to over 10,000 years at 70 °C, which is clearly excessive. With the release of JESD-49I in 2012, a separate test condition was specified for PCM, with the bake temperature reduced to 90 °C [8]. With 1,000 hours of bake stress, this becomes equivalent to the floating gate expectation of 10 years at 70 °C. Even if this condition is very achievable with GeSbTe-based PCM devices, specialized applications requiring continuous, very-high-temperature operation (e.g., greater than 100 °C) should increase the bake retention qualification temperature as required to capture the specific usage condition.

As mentioned at the beginning of this section, in single-level cell implementations of PCM, RESET crystallization is the main data retention concern as it brings the RESET state closer to the SET state resistance. The other mechanism of resistance shift with time, drift, also impacts the RESET state but increases the cell resistance (and thereby the window between the states), so it is not a concern. While multilevel cell (MLC) PCM will not be discussed here in detail, it is worth noting two key differences in data retention from the single-level cell. (A more detailed study of MLC PCM can be found in Papandreou et al. [9].) When considering extending PCM to MLC, cells must be placed in – and remain in – states between the SET and the RESET state. This has two ramifications. First, the partially RESET state will drift to higher resistance with time, and adequate resistance window must be available to accommodate this increase. And second, since the intermediate states most likely have been programmed with a “soft” RESET pulse, it may crystallize faster than when programmed to a fully RESET state (as was shown earlier in this chapter). Since these mechanisms shift the resistance in opposite directions (first increasing, then declining), the worst-case condition for cell failures may be at shorter times and/or temperatures than in the case of single-level cell PCM. In this case a full assessment of cell failure rates across a range of times and temperatures should be performed to assess the correct conditions for product qualification.

5.4 Cycling Endurance

One of the primary advantages of PCM over flash memory is the reduced write latency. This is enabled by the ability to arbitrarily change data in the PCM array at the address level (as opposed to erase at the block level). Two consequences of this capability are increases in both the maximum writes that are possible on a memory cell and the variation in cell-to-cell writes. As a result, the write requirement for PCM is several orders of magnitude higher than flash and is typically quoted as 10^6–10^8. In fact, this requirement could easily be exceeded in a product. With typically write times near 1 μs, in the extreme case (continuously writing), a single cell could be written 10^{14} times over the course of 10 years. While this is not a likely condition, it does suggest that PCM specifications are somewhat less than what would be required to remove all concerns about write endurance. (And therefore a careful analysis of the write patterns or a method to prevent the extreme case is warranted.)

Early write endurance studies of PCM showed a very high intrinsic cell capability, often $>10^{10}$ writes. An example of this is shown in Fig. 5.10a, where data were collected on a single PCM cell [10]. In this example, the cell showed very stable RESET and SET resistance throughout the lifetime, becoming unstable at ~10^9 cycles until (presumably, since it is not clear from the figure) the cell failed at 10^{11} cycles. In the same reference, PCM cell arrays were cycled to show the impact of statistical variation on cycling capability. In this case the resistance of the cells at various standard normal quantile (σ) values was reported for both the SET and RESET states. In this case we see the first loss of cell window to occur from the SET side, where the -4σ (3×10^{-5}) cell shows a steep decline at 10^7 cycles.

When considering the cycling capability of a PCM device, it is important to distinguish between the mechanisms that are intrinsic to the chalcogenide material and those that are dependent on the specific surrounding materials and integration scheme. Writing a PCM device exposes these materials to both high temperatures (>650 °C) and, for the case of the chalcogenide and electrodes, high current densities (10^9–10^{10} MA/cm^2). At these conditions, the material has both a high atomic mobility and a driving force for possible electromigration-driven transport. Studies have been performed on simple cell structures to examine these driving forces and the resultant (intrinsic) material transport. Yang et al. performed experiments on a line structure to understand the elemental separation that occurs when subject to current pulses that place the cell in the molten state (above melting temperature) or crystalline state (below melting temperature) [11]. Figure 5.11a and b show the device that was used to perform these experiments and the electrical current waveforms that were applied. The results for the molten state are shown in Fig. 5.12. In this case the cell was stressed to the point of breakdown, though at intermediate states, the composition showed an increasing level of separation with increasing stress time. These data show that in the melt state the Te takes the form of an anion, while the Ge and Sb act as cations, moving in opposite directions. Furthermore, it was reported that doping the line with nitrogen could slow the rate

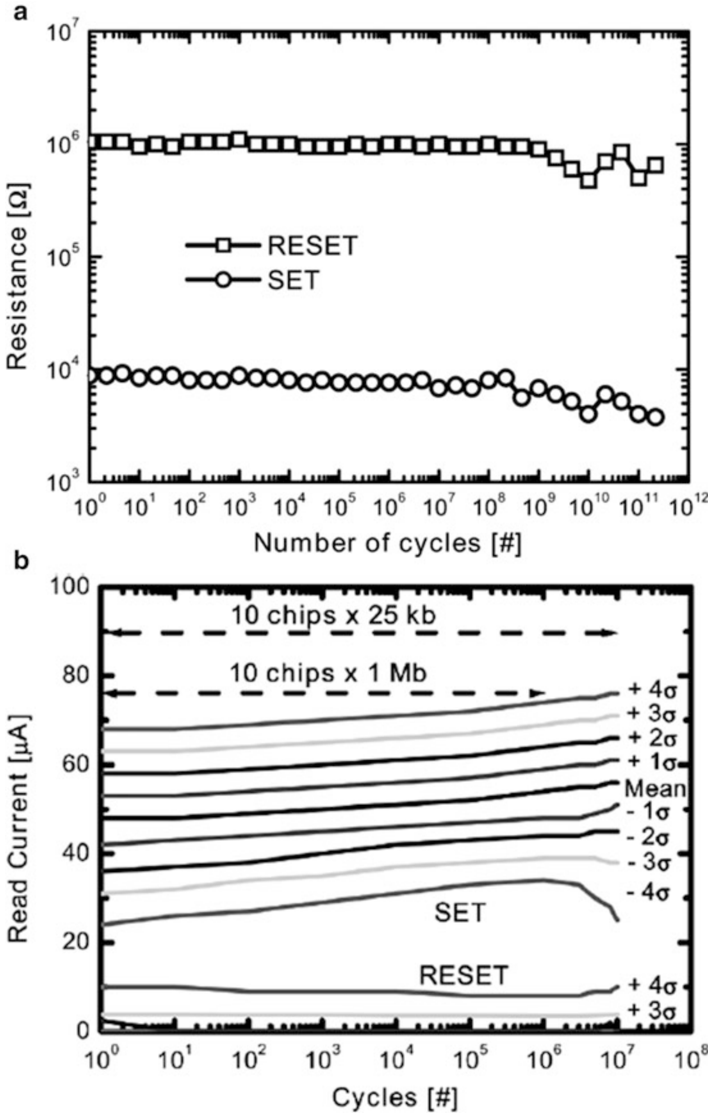


Fig. 5.10 (**a**) shows greater than 10^{11} intrinsic endurance capability of a PCM cell. (**b**) shows that, on a distribution of many cells, the write endurance on the earliest failing cells (here the -4σ normal quantile) can be significantly lower (From Pirovano et al. [10])

of migration. When stressing in the crystalline state (Fig. 5.13), the amount of compositional variation was much lower and in the reverse direction. Comparing this with an unstressed device, it appeared that all the elements were moving toward the cathode and that the composition difference was dependent upon the relative speed of migration.

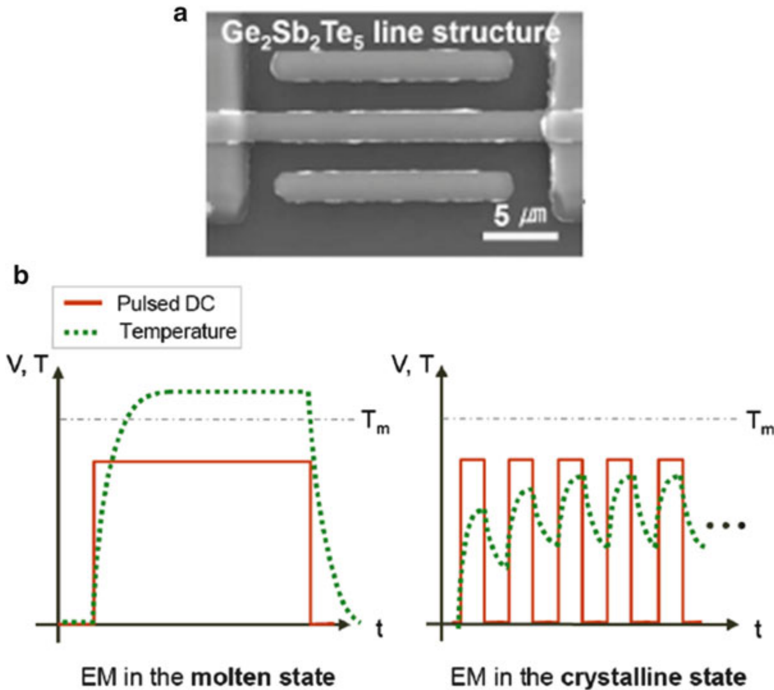


Fig. 5.11 (**a**) shows the device used by Yang et al. to investigate elemental migration in $Ge_2Sb_2Te_5$. (**b**) shows the voltage waveforms that were applied to the cell to stress it in either the liquid or crystalline state [11]

The relative importance of cycling with low current pulses (i.e., in the SET state) vs. higher current (the RESET state) was studied by Kim and Ahn using an N-doped GST cell with a 60 nm electrode contact diameter [12]. Here they found that relatively low current pulses (0.6 mA) did not result in cell failures to 10^{12} writes, which was the testing limit. On the other hand, with a RESET pulse of 1.2 mA applied, resistance of the RESET state began to decline after 10^9 writes. Given this degradation, further studies were run to understand the acceleration of failures with programming conditions. In a large memory array, providing low write latency requires the programming current to be set at a level where most of the cells are programmed in a single write operation. (Of course an adaptive, multistep programming algorithm can relax this requirement, but typically increases both write energy and latency.) As a result, a significant number of cells must be programmed with a higher current than their optimal level. Kim and Ahn found that the cycles required for failure is inversely proportional to the energy of the programming pulse (as shown in Fig. 5.14a) [12]. Cross sections of over-reset cells show a significantly larger amorphous region than is seen in the optimally programmed cells (Fig. 5.14b). An analysis of write endurance failure modes is shown in Fig. 5.15. Figure 5.15a shows the electrical characteristics of failing cells, which indicates

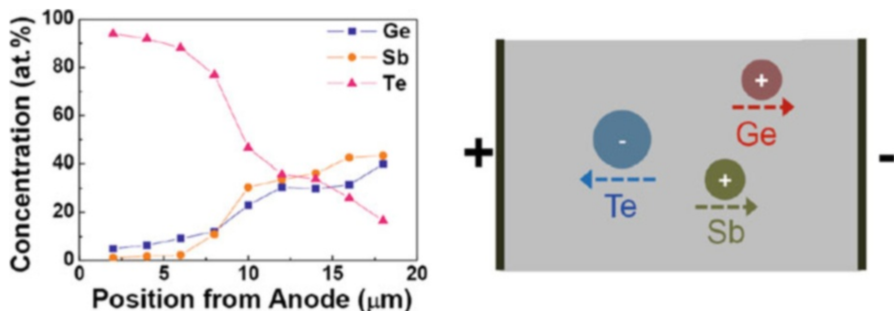


Fig. 5.12 Elemental segregation as observed in the line device structure when stressed in the liquid state [11]

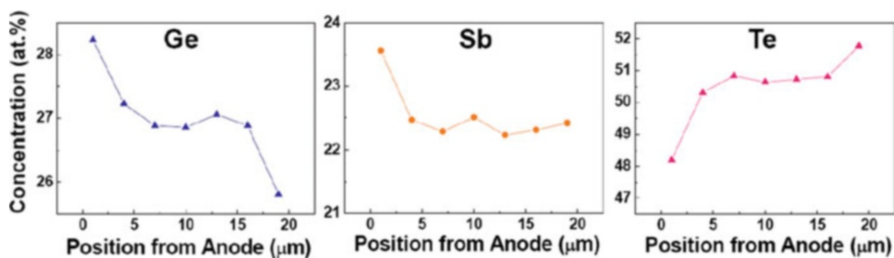


Fig. 5.13 Elemental segregation as observed in the line structure when stressed in the crystalline state [11]

that the stuck SET fails primarily come from the initial over-reset distribution. Figures 5.15b and c contain micrographs of failing cells. These indicate that the stuck SET cells are the result of Ge depletion in the programmed material adjacent to the heater interface, while stuck RESET cells are the result of void formation at the electrode interface.

While the above analyses concluded that the RESET operation is responsible for cycling-induced damage, other reports suggest the behavior is more complex. Du et al. examined the impact of varying the programming current on the cycling endurance of a mushroom cell [13]. In their experiment, the nominal current required to RESET the cell was in the range of 200–300 μA. Their programming waveforms for RESET and SET typically started above the melting current and differed in the waveform timing: the rising edge/plateau/ramp-down stages were 19 ns/100 ns/10 ns for RESET and 100 ns/300 ns/1000 ns for SET. As a result, and unlike the analysis of Yang et al. discussed above, the GST is melted during the application of the SET pulse as well as the RESET. With this approach, Du et al. found that (for an optimized RESET current) cycling capability decreased from 10^8 to 10^6 as the current of the SET pulse was increased from 220 μA to 350 μA. (The results are shown in Fig. 5.16a.) Increasing the current of the RESET pulses,

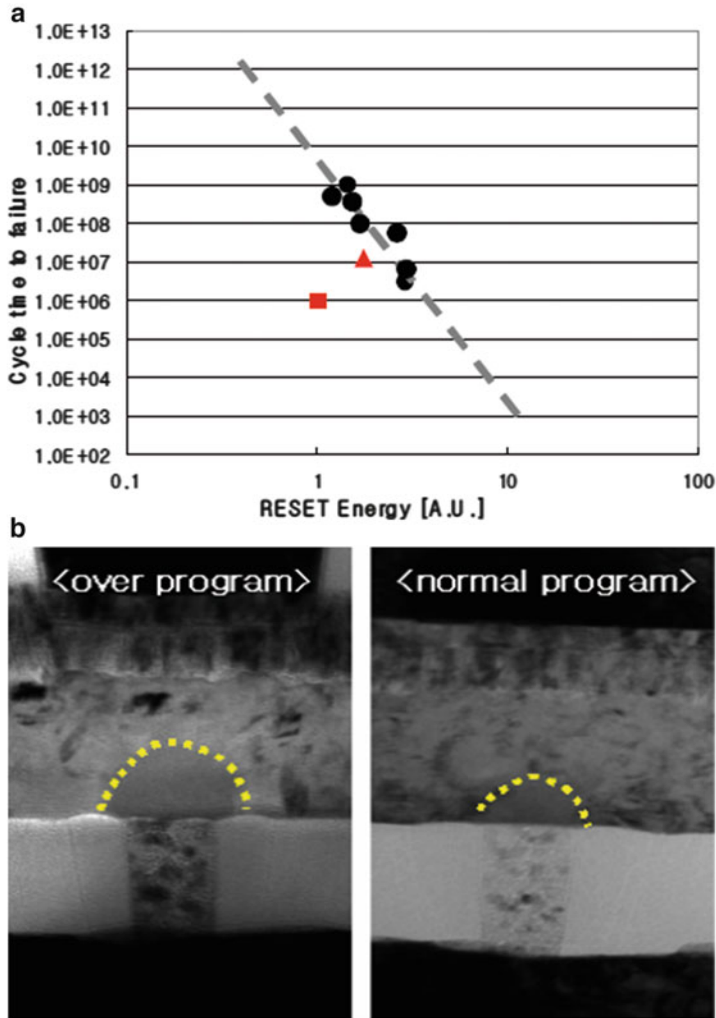


Fig. 5.14 (**a**) shows the decline in write cycles to failure with the energy of the applied RESET pulse. (**b**) contains micrographs of cells that were RESET with normal and overprogrammed levels [12]

however, showed the opposite trend (Fig. 5.16b). In this case the current was varied from 190 to 420 μA, and the median cycling capability increased from 10^6 to 10^8.

The above analyses suggest that write endurance in PCM is limited by current-driven (and temperature accelerated) migration of elements. Whether the cell fails due to a stuck SET or stuck RESET is determined by the dominant mechanism: compositional segregation limiting the ability to amorphize the programmed region or material depletion resulting in a void or delamination at the interface. To further understand the impact of current on PCM material, Lee et al. studied the impact of

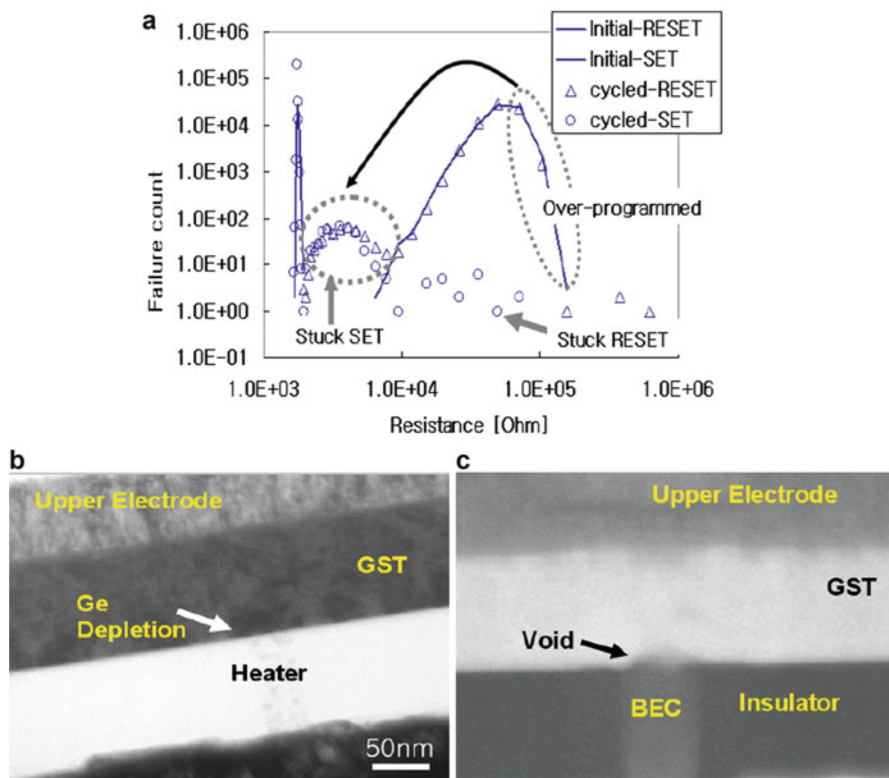


Fig. 5.15 (**a**) shows the cell counts as a function of resistance for both the initial and post-cycling condition. (**b**) and (**c**) show the mechanisms behind stuck SET and stuck RESET failures as Ge depletion and voiding at the electrode interface, respectively [12]

current direction on cycling endurance [14]. In their experiments, cycling was performed both in the "forward" (chalcogenide-biased positive vs. the electrode contact) and "reverse" directions (chalcogenide-biased negative). In addition, they ran experiments where the bias condition was varied between the RESET and SET operations during cycling. Electrical behavior from these cycling variants is shown in Fig. 5.17. These results show that the cell is clearly nonsymmetric with regard to the programming current direction, with forward-SET + forward-RESET providing both the largest window and the highest write endurance. In contrast, reversing the SET current showed a complete lack of window and very noisy programming behavior, while moving just the RESET pulse to the reverse direction showed some (though degraded) cell functionality. Figure 5.18 shows the composition variation through the programming region for the two extreme cases. In the top figure (forward-forward programming), Sb enrichment is observed near the interface, but otherwise the material is intact. On the other hand, with reverse-reverse current cycling, bulk material depletion is observed, and failure takes the form of a void at the interface. As with the prior results, these studies confirm the fact that the

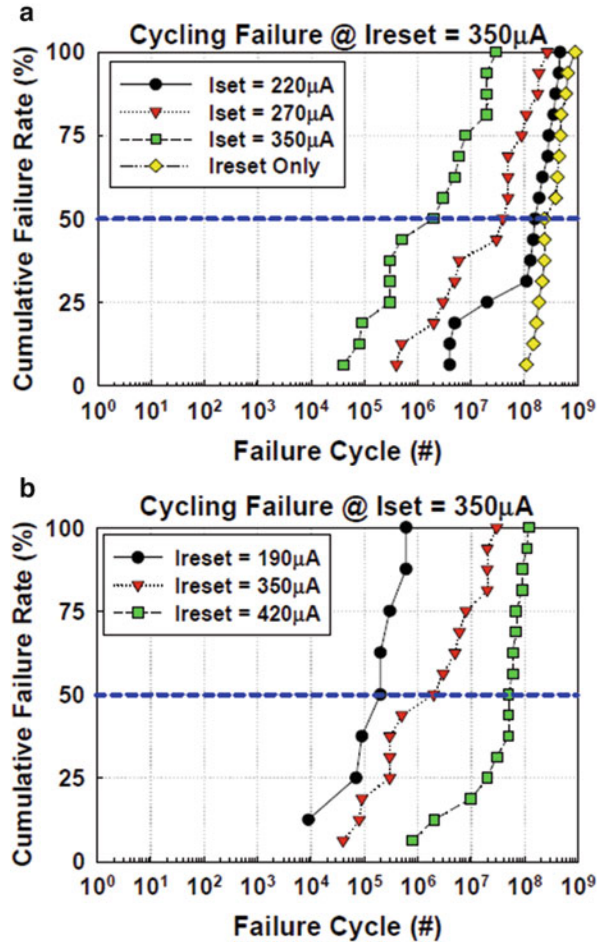


Fig. 5.16 (**a**) Cycling failure distribution using (**a**) SET-only and (**b**) RESET-only cycling (From Du et al. [13])

write endurance fails are not simply temperature driven (by joule heating) but instead heavily dependent on a current-induced transport. An extensive discussion on ion transport modeling is provided in Chap. 3.

Given the material segregation during write cycling is dependent upon the programming pulse conditions, several methods have been proposed that use a modified write pulse to "heal" the damage induced by write cycling. Such a process is claimed in US Patent 8,036,016, where a modified pulse is used to reverse the damage [15]. In this case, the memory subsystem would include the ability to detect the accumulation of damage to cells (by counting write operations, reacting to an increase in bit error rate, or other means). When a specific threshold condition is met, the healing pulse is applied. While broad in its claims, the patent describes a pulse that is different from the typical programming pulses (and is typically higher in current or longer in time). In recent study, Khwa et al. described a specific implementation of this concept with positive results [16]. After observing that a

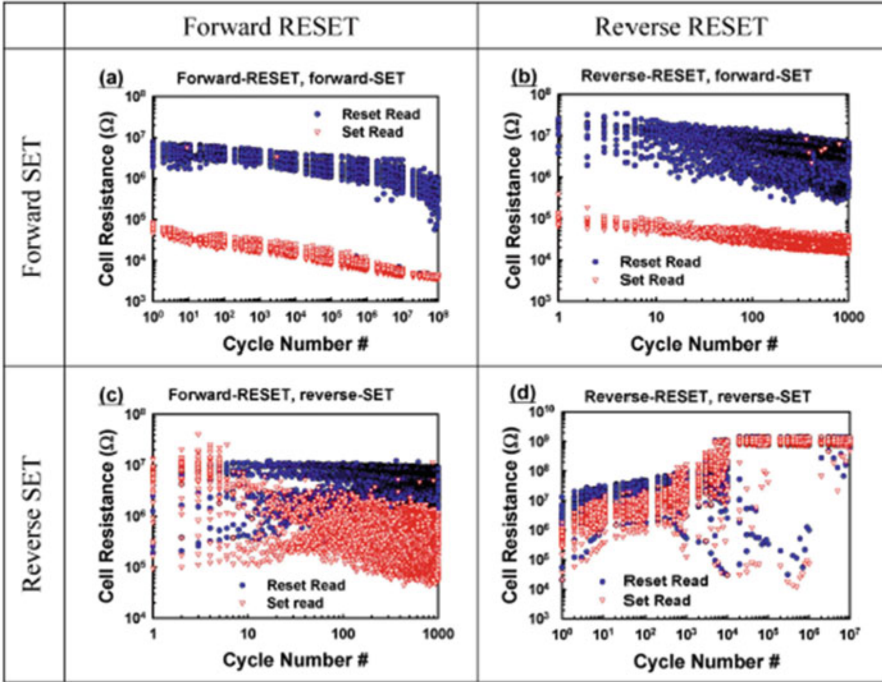


Fig. 5.17 Impact of varying current direction during the RESET and SET operations in a PCM cell (From Lee et al. [14])

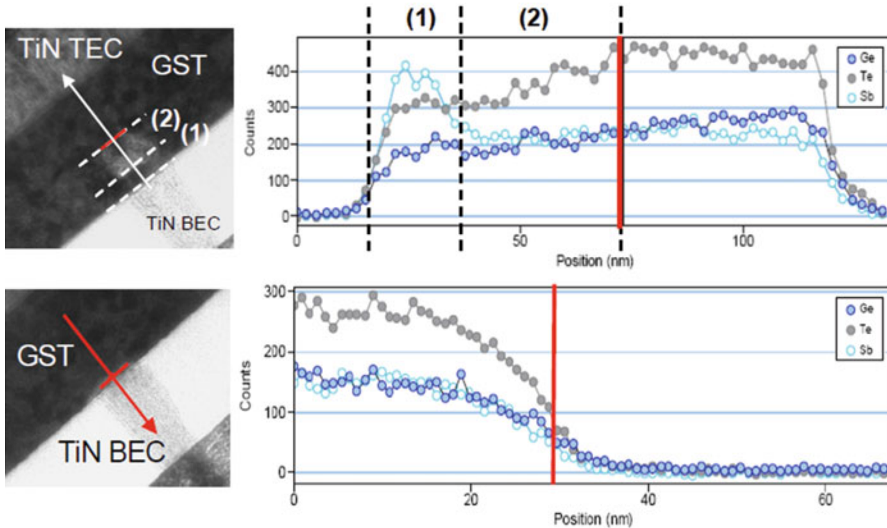


Fig. 5.18 Chalcogenide composition vs. distance from the electrode interface after forward-RESET + forward-SET cycling (*top*) and reverse-RESET + reverse-SET cycling (*bottom*) [14]

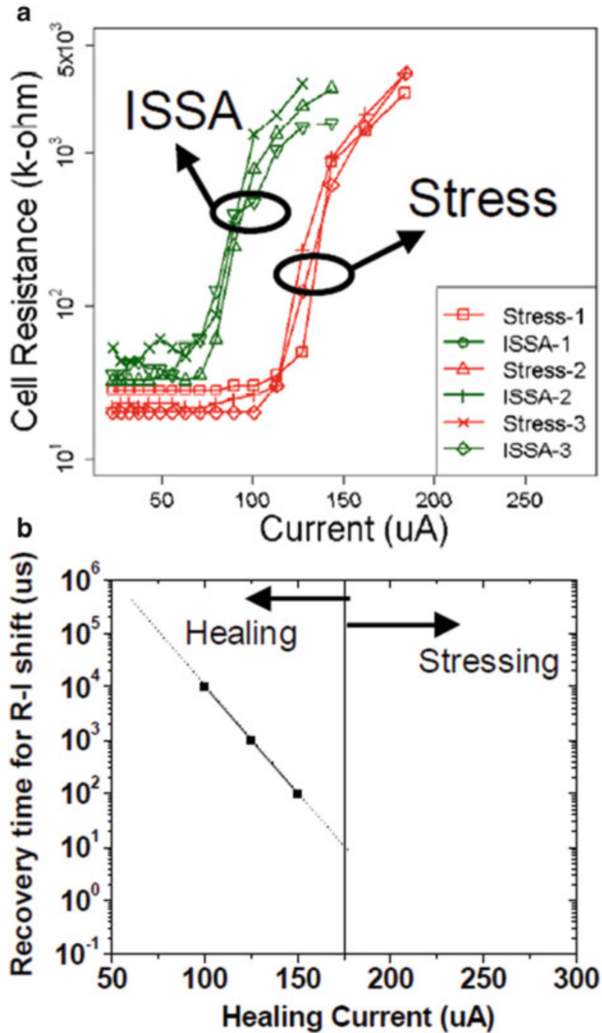


Fig. 5.19 (**a**) shows the R-I curves of a cell that was subjected to repeated cycling stress and a subsequent ISSA (in situ self-anneal) recovery pulse. (**b**) shows the dependence of recovery time on healing current, indicating that the recovery time is faster at higher currents to a critical level, above which the pulse conditions create additional damage to the cell [16]

350 °C anneal could restore the programming characteristics of a heavily cycled cell to its initial state (and then subsequently allow it to cycle normally), they attempted to replace the bake with an electrical waveform with the idea that electrical heating could achieve the same result. They observed that this is possible, but only if the cell damage is recovered before it becomes too severe. After cycling, the R-I curve for the cell (resistance vs. the amplitude of a square programming pulse) moved to the "right," indicating an increase in the current required to RESET the cell. This shift was monitored by applying a low-current programming pulse. Once a significant decline in the resistance was measured, the recovery pulse was applied. Figure 5.19a shows this behavior, where the green lines correspond to the initial and "recovered" R-I curves and the red curve corresponds to the post-cycling

stressed curves. As this figure indicates, the cell can be recovered from the cycling-induced degradation multiple times. Also discussed in the Khwa paper was the importance of the recovery pulse conditions (both current and time). Figure 5.19b shows the recovery pulse time as a function of the healing current, where higher current can reduce the time. Beyond a critical current, however, the recovery effect was not seen, and instead the cell was exposed to additional damage.

5.5 Read Disturb

Similar to data retention, read disturb is a failure mode with crystallization of the amorphous PCM phase as the underlying mechanism. In this case, however, the crystallization is accelerated by both the ambient temperature and the applied bias of the reading operation. (In fact the mechanism is more likely driven by currently flowing through the cell during the read, even if the bias is the controlled parameter.) While the bias acceleration can be significant, since the voltage is only applied during the reading operation, the total time in this state is significantly reduced when compared to the requirement for standard (unbiased) retention. Table 5.2 shows the total bias time required for reading over a 10-year product life as a function of the array size (with the assumption that the bias is applied for 100 ns during the read operation, typical for PCM). This table also includes the sensitivity to array size and how the reads are distributed over the array with time. One extreme condition for reading is that where the entire array is read continuously but uniformly (the "Uniform reads" column). In this case the required number of reads depends on the density of the component, with the read requirement ranging from 10^4 to 10^7 on products from 100 Gb to 100 Mb. At the other extreme, we consider the case where a single cell is read continuously over the lifetime of the product. Here the number of required reads is very high, near 10^{15}. While it would clearly be an ideal situation for the PCM device to be able to sustain this number of reads, the usage case is very unlikely, and there are means to manage this at the product level (e.g., by implementing a read cache or a mechanism to refresh). Perhaps a more likely scenario is the case where a small portion of the array is read very aggressively while the remainder is accessed at a much lower level. This is considered in the "1% array reads" column, where we consider a case where 1% of the array is subject to frequent reading. In this case the read requirement is between 10^6 and 10^9. While the read requirement for a product must be consistent with the

Table 5.2 Number of reads sustained on a PCM cell over a 10-year product usage, considering different array sizes and the distribution of reads

Product size	Uniform reads	Worst-case reads	1% array reads
100 Mb	10^7	10^{15}	10^9
1Gb	10^6	10^{15}	10^8
100Gb	10^4	10^{15}	10^6

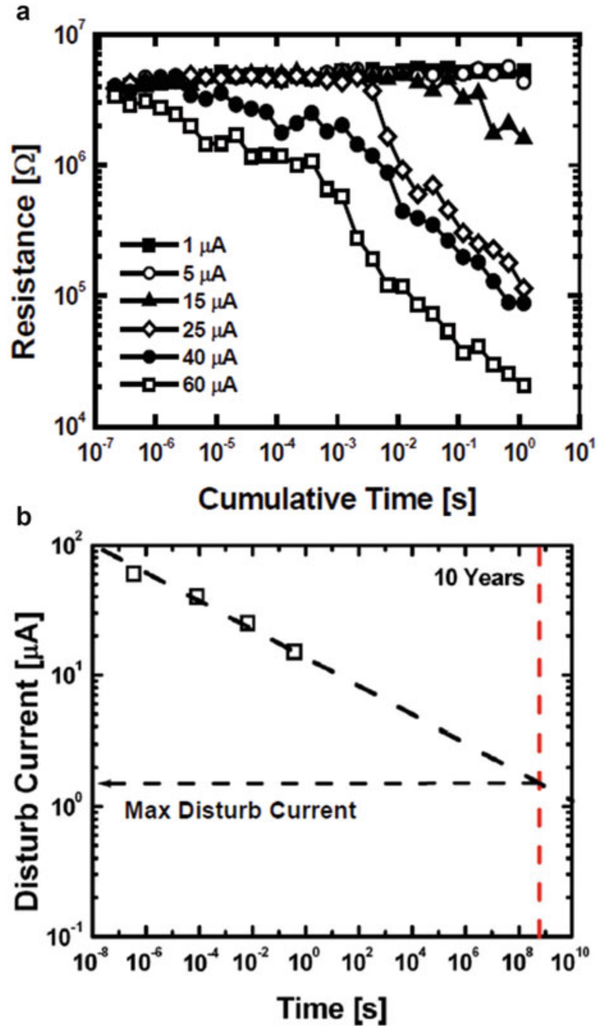


Fig. 5.20 (**a**) Resistance decrease with time as a function of applied current and (**b**) the extrapolated reading current that could be applied for 10 years of continuous reading (Both figures are from Pirovano et al. [10])

expected application, from this assessment we could consider 10^{10} reads as a reasonable worst-case requirement.

An early examination of read disturb in PCM was performed by Pirovano et al. [10], who measured the resistance change with time while applying a variety of fixed currents. Their results are shown in Fig. 5.20a, where the time-to-disturb decreases with increasing applied current. The mechanism for this disturb was hypothesized to be a very localized heating of the cell in the amorphous phase, as the low applied current was insufficient to produce significant uniform heating across the cell. In this case the device did not show signs of disturb to 1 s of total stress time once the current was reduced to 5 uA, which would correspond to about 10^7 read operations (with 100 ns of bias during each read). To estimate the read

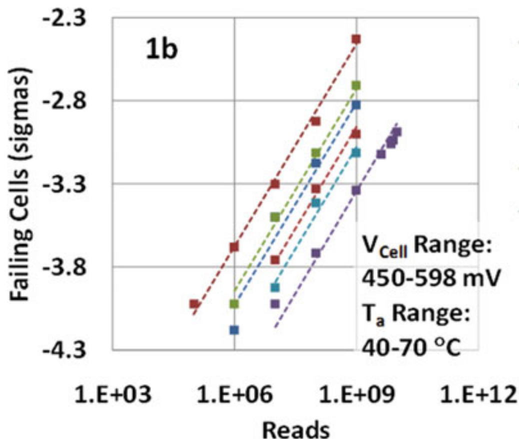


Fig. 5.21 Read disturb cell failure rate vs. number of reads at various voltage and ambient temperature conditions. Increasing either the voltage or the temperature shifts the curve to the *left*

current required for a full 10-year continuous read capability (the extreme case in Table 5.2), the time-to-failure was estimated by a 2x reduction in cell resistance, and the current vs. time curve was extrapolated to 10 years. This is shown in Fig. 5.20b and indicates that the required current through the RESET state during read should be less than 1 μA. This is about ten times higher than the typical read current, which should provide a very high intrinsic capability.

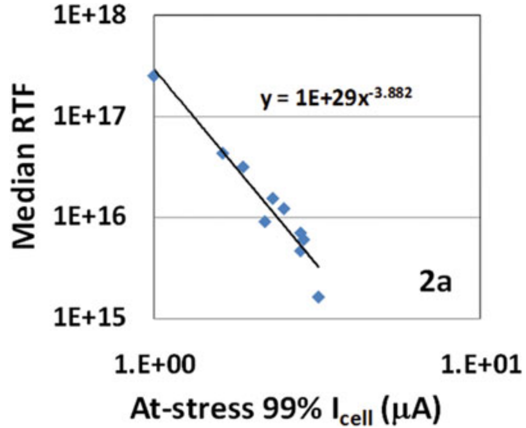
A more recent study of read disturb characterization was performed on a commercial PCM product [17]. The device was read across a range of voltages (the controlled variable in this case) and ambient temperatures, and the rate of cell disturb was measured. Figure 5.21 shows the fraction of cells disturbing (in units of "sigmas" from a standard normal distribution) with increasing number of reads. Each line represents a combination of read voltage and ambient temperature, with the curves moving to the left as either temperature or voltage is increased.

While these data show a dependence of read disturb on both voltage and temperature, it is important to understand whether they directly contribute to disturb or indirectly contribute through their impact on the reading current. Several models have been proposed in the literature to connect these variables, one being the model derived by Zhang and Ielmini that has shown a good correlation to experimental data [18]. In this case the subthreshold current is given by

$$\mathrm{I}_{\mathrm{cell}} = c\sinh\left(\frac{AV_{\mathrm{cell}}}{kT_{\mathrm{a}}}\right)\exp\left(\frac{-E_{\mathrm{a}}}{kT_{\mathrm{a}}}\right),$$

where V_{cell} is the applied voltage, T_{a} is the ambient temperature, and the remaining parameters are constants. Fitting the read disturb data to the read current (i.e., the current flowing through the cell at the applied read voltage and ambient temperature), we find that the median reads to failure is dependent only on this current, regardless of the ambient temperature or the applied voltage. This is shown in Fig. 5.22, where the median reads to failure as extracted from 5.22 is plotted against

Fig. 5.22 Median reads to failure (extrapolated) as a function of the reading current through the cell



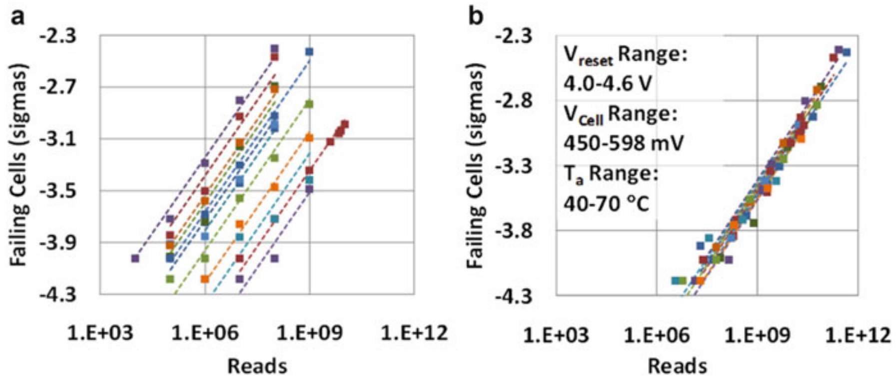


Fig. 5.23 Read disturb failure rate at different applied temperature, voltage, and initial RESET voltage conditions. (**a**) contains the unscaled result, and (**b**) contains the results where the reads were scaled by ${I_{cell}}^{-4}$

the cell read current. These data indicate a power-law dependence with an exponent of approximately −4.

While these data suggest that the reading current is the primary factor governing read disturb, additional experiments were run that include both the voltage and ambient temperature as above, but also the RESET programming condition. To check a more general case, read disturb was measured in an experiment where the read current was varied by modulating the read voltage and ambient temperature (as above) and additionally the RESET state itself (by varying the programming condition). The results are shown in Fig. 5.23, with the actual data plotted in 5.23a and the distributions normalized by scaling the reads by the factor ${I_{cell}}^{-4}$. These data again suggest that the read disturb capability is dependent only on the reading current. It also indicates that, if the reading current is kept to a level less than 1 uA

(which is typical for PCM), the read disturb capability is more than adequate to achieve the required product specifications.

5.6 Thermal Disturb

In additional to read disturb, a second disturb mode in PCM devices can result from thermal cross talk between neighboring cells during write operations. Similar to read disturb, the underlying mechanism is an undesired crystallization of a RESET cell. But unlike read disturb, thermal disturb results from heat flow between a cell that is being programmed and a neighboring cell (known as the aggressor and victim cells, respectively). Since the temperature, and thereby the risk of disturb, is highest in cells that are closest to the aggressor, thermal disturb is often called proximity disturb.

The thermal disturb requirements for a PCM array depend largely on the expected writing patterns of the application. On one extreme, we would consider a cell that is programmed at the time the component is put into service with no subsequent reprogramming. After this initial programming, all of the neighboring cells are programmed to the full write endurance specification of the component. In this case the requirements could be as much as four times the write endurance specification. This assumes that only immediate neighbor aggressors provide significant heating, and they are all equal in terms of impact to the victim. (While the first assumption is generally valid, the second is often not the case.) With this condition in mind, the thermal disturb tolerance would need to be in the range of 10^7 –10^9 writes. A second extreme case would occur if the component was used in a similar fashion to flash memory, where cells are programmed and erased in a larger "block"-like fashion. Here a victim cell would only be subject to a single cycle from the aggressors, and therefore the thermal disturb requirement would be only a few writes. While it is possible to design a system that manages aggressor writes to any level between these extremes, the level of complexity and performance impact will increase as the allowed aggressor writes is reduced. If we consider a reasonable case where the aggressor writes are limited to 1% of the cycling specification (at which point the victim must be refreshed), the required thermal disturb capability would be in the range of 10^5–10^7 writes.

Initial assessments of thermal disturb risk were performed using thermal simulations, as is shown in Fig. 5.24 [19]. When this paper was presented in 2003, the cell spacing in PCM test arrays was typically greater than 100 nm. Since the main interest was to develop products with a scaling path to less than 50 nm separation, modeling was used to estimate the temperature as a function of distance away from a cell being programmed (which would correspond to the temperature at the victim cell). In this case the chalcogenide material was assumed to be at its melting temperature (~650 °C), and the temperature was modeled as a function of distance and programming pulse widths. The results show the rapid decline in temperature with distance, to the extent that the temperature is less than 100 °C at a distance of

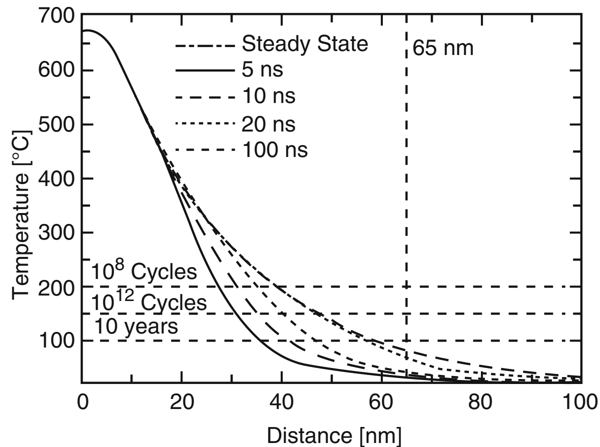


Fig. 5.24 Simulations of temperature as a function of distance from a cell during a RESET operation [19]

65 nm from the aggressor cell. Moreover, the pulse width analysis indicates that at times below 20 ns, which are typical for a RESET pulse, the thermal latency of the cell leads to a significant reduction in the victim temperature vs. the temperature that would be reached in steady state. Based on these analysis and using data retention experiments to estimate the time to crystallization, the expectation was that thermal disturb should not be a significant concern at 65 nm spacing. Furthermore, with very short RESET pulses, the cells should be able to tolerate 10^9 cycles without disturb at less than 40 nm spacing.

To better predict the impact of scaling on thermal disturb, a more detailed analysis was performed by Russo et al. [20]. This analysis considered the impact of scaling not just the dimensions but also the aspect ratio of the cell. While this analysis was performed on two different cell architecture ("lance" and "wall") cells, the concepts were similar, and only the lance cell analysis will be shown here. Figure 5.25 shows a 2-D cross section of the cell with dimensions corresponding to a 90 nm technology node. Here the two vertical elements are heating electrodes that are connected to a line of chalcogenide above them, and the active phase change region is in a "mushroom" shape immediately above the electrode. Two memory cells are included in the model, with the cell on the left acting as the aggressor and the cell on the right is the victim. Thermal simulations were performed on this structure both for the nominal 90 nm structure (a) and with the cell scaled in two different ways. In case (b), the lateral dimension of the cell was scaled to 45 nm, while the vertical dimension was kept constant (non-isotropic scaling). In case (c), all dimensions were scaled uniformly (isotropic scaling). The overlaid thermal gradient lines reflect the temperature during programming with red reflecting hotter and blue as cooler temperatures. As the figures indicate, the victim temperature will be significantly higher in the case of non-isotropic scaling. This temperature is quantified in Fig. 5.26 (also from Russo et al. [20]), where the victim temperature is simulated as a function of scaling dimension and for the case of isotropic and non-isotropic scaling. (A third, intermediate scaling condition is shown here as well

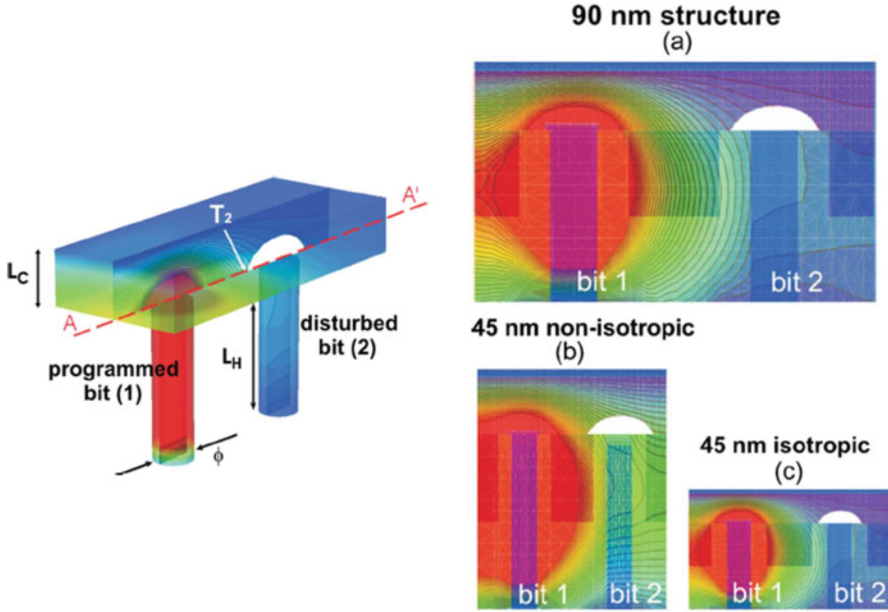


Fig. 5.25 Cell structure modeled by Russo et al. to examine the impact of aspect ratio scaling on thermal disturb [20]. The figure to the left contains the modeled cell structure, with a chalcogenide line above two vertical (and heating) electrodes. The figures to the right include (**a**) the thermal gradient lines at 90nm, and in (**b**) and (**c**) the *gradient lines* when scaling just the lateral dimension (**b**) or all dimensions (**c**) to 45nm. In case (**b**) the maximum victim temperature is significantly increased with scaling, while in (**c**) it is constant with scaling

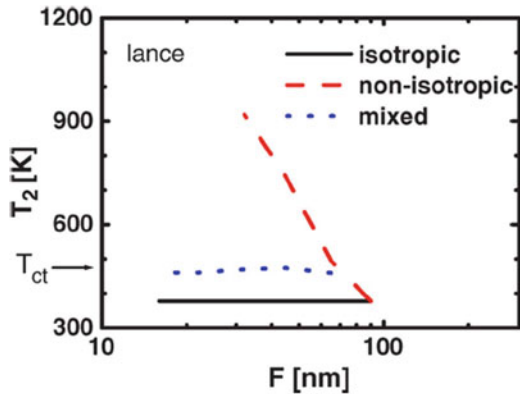


Fig. 5.26 Quantification of the victim cell temperature (T_2) vs. cell dimensional scaling (F) as described in Fig. 5.25 [20]

and discussed in the original paper.) As this analysis indicates, in the case of isotropic scaling, the victim temperature remains constant, and the thermal disturb capability should not be reduced.

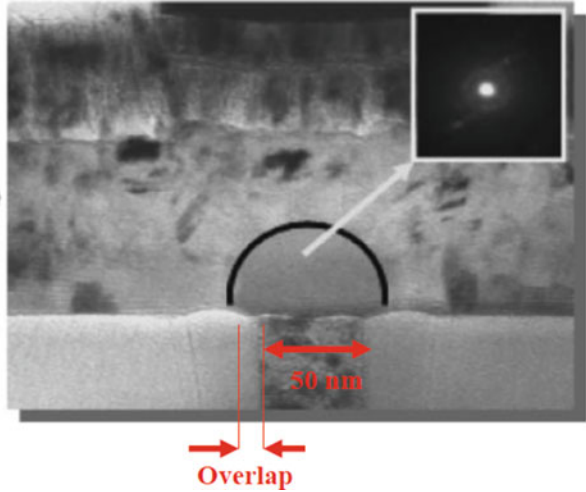


Fig. 5.27 The "overlap" region in the lance PCM cell, as defined as the required distance the amorphous region must extend beyond the electrode heater to provide the required RESET resistance (From Hudgens [21])

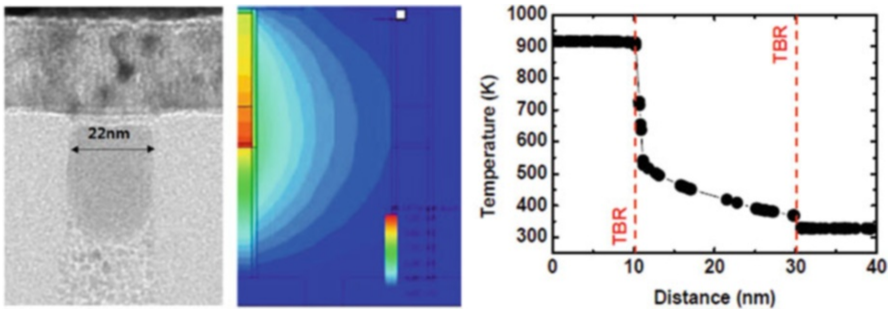


Fig. 5.28 Temperature as function of distance with thermal boundary resistance [22]

While the prior analysis shows a high potential for scalability, a concern was raised by Hudgens as to whether fully isotropic scaling is possible in cell structures where the amorphous region is "unconfined" [21]. Specifically, for the lance cell to achieve a high resistance, an "overlap" region must be maintained to isolate the electrode from the bulk chalcogenide (as shown in Fig. 5.27). Since this overlap requirement is based on the amorphous chalcogenide resistivity, it is not expected to scale down significantly with cell dimension. As a result, the edge of the amorphous region would move closer to the aggressor cell as the spacing is reduced, increasing the risk of disturb. While this dimension may be a relatively small component of the cell separation at 40 nm spacing, it suggests that an alternative cell architecture could be required to enable further cell scaling.

One direction toward a more scalable cell is to confine the chalcogenide in a small area, with electrodes on both sides (as is shown in Fig. 5.28, from Kang et al. [22]). While this cell architecture has generally been pursued as a path to

programming current reduction, the analysis of Kang et al. shows the potential for reducing thermal cross talk. In addition to removing the overlap region required in an unconfined cell configuration, the confined cell has the additional benefit of introducing one or more material interfaces between the cells. The resulting thermal boundary resistance (TBR) at these interfaces has the effect of reducing the thermal transfer between the cells. From the simulations show in Fig. 5.28, the impact of the TBR could be equal to or larger than the thermal impedance of the bulk dielectric films between the cells. Moreover, since the TBR is not reduced with scaling, any advantage it provides is not reduced with further reduction in cell spacing. From these simulations and comparing to reference GST crystallization data, the 20 nm confined cell is expected to tolerate 10^7 aggressor writes before disturb.

While the prior analyses show the potential to maintain intrinsic thermal disturb resistance with cell scaling, the statistical variation of disturb across a large array must also be considered. This variation typically arises from three sources. First, to program a robust RESET state into a large number of cells, the current must be chosen based on the cells with the highest requirement. As a result, most of the aggressor cells receive more current than is required to melt and therefore are heated beyond the melting temperature (over-reset). Second, the dimensions between cells will have a statistical variation, with the closest aggressor-victim pairs having higher risk of disturb. And third, as discussed in the earlier section on data retention, we expect there to be a variation in the crystallization rate of the victim cells at a given disturb temperature. Here the fastest crystallizing cells have the highest risk. To fully quantify the relative impact of these elements would require an understanding of the relative spread of each as well as the impact to thermal disturb.

A statistical level of thermal disturb for a scaled (45 nm) PCM technology was reported by Redaelli et al. [23]. The cell structure used in this study is shown in Fig. 5.29, which shows both the cell and array architecture as well as the crystallization of a disturbed victim cell. In this architecture, the primary direction of disturb is along the bitline, where the aggressor and victim cells are connected through the continuous chalcogenide line. Disturb in the lateral direction, which includes the inter-metal dielectric as well as the associated interfaces, was very

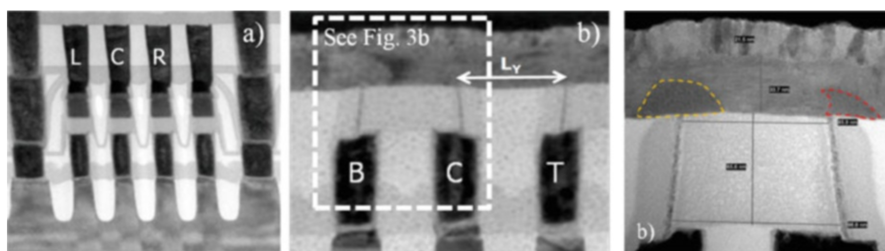


Fig. 5.29 Cross sections of the 45 nm PCM array used in the study of Redaelli et al. [23]. The picture to the *left* shows the geometry orthogonal to the bitline (not sensitive to disturb), while the center picture shows the cell separation along the bitline (the primary disturb direction). The picture to the *right* shows the partially crystallized amorphous region in the victim cell (outlined with the *dashed red line*)

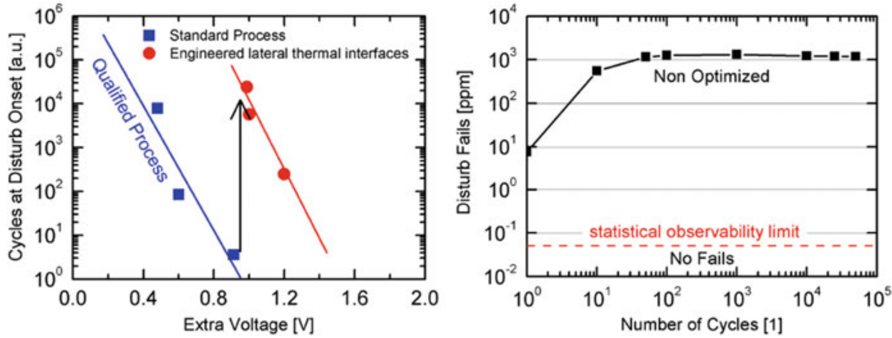


Fig. 5.30 Graphic on the left shows the relative disturb with extra voltage (which corresponds to over-reset of the aggressors) [23]. The graphic on the right shows the cell failure rate with aggressor cycles at the standard programming conditions. These shows the results before and after cell process optimization, which is able to bring the disturb rates below the required level

robust. The results of this study are shown in Fig. 5.30. Specifically, while the sensitivity to over-reset is quite large, with proper optimization of the cell parameters (including reduction in variability in the cell spacing), the 45 nm array shows the capability to exceed 10^5 writes without a detectable disturb.

5.7 Radiation Sensitivity/Soft Errors

Since the storage mechanism in PCM is not based on electrons, it is not expected to share the radiation sensitivities of floating gate memories such as NAND or NOR flash. In these technologies, impinging particles create electron-hole pairs that can shift the apparent threshold voltage of the cell and result in soft data errors. A PCM device should be much more resilient to this mechanism. In addition, the demarcation currents for PCM are in the μA range and therefore much higher than what would be locally created by radiation exposure. While published data are fairly limited, particularly on scaled PCM cells, they indicate that while the bulk of the chalcogenide material is robust, there is the potential to disturb at the heater interface in a highly scaled cell.

In one of the earliest studies of radiation hardness of PCM, test structures were developed based on a radiation-hardened CMOS fabrication process by Maimon et al. [24]. These arrays were subjected to both total dose testing to 2 MRad and single event upset (SEU) testing to a maximum effective linear energy transfer (LET) of 132 MeV/mg/cm^2. The total dose testing showed no shifts in the electrical characteristics of the PCM memory cell. While the SEU testing did expose both a false read and a read disturb mechanism, the source of failure was isolated to the CMOS circuitry on the test array, and the lack of any failure in the PCM element was confirmed.

More recently, heavy ion-induced SEU testing was performed on a 1 Gbit PCM device based on a 45 nm process technology from Micron Technology [25]. The

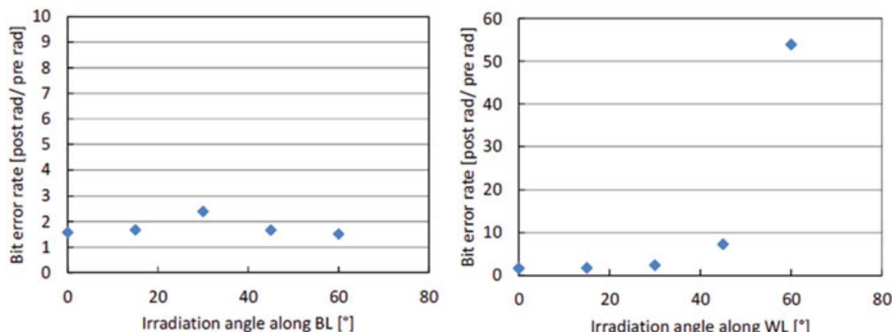


Fig. 5.31 BER before and after heavy ion exposure with the angle of incidence along the bitline direction (*left*) and along the wordline direction (*right*). Note that BER ratios between 0 and 2 are considered to be within the background noise [25]

cell in this device incorporated a thin heating element contacting the chalcogenide, the so-called "wall" architecture [26]. Two types of testing were performed: wide energy spectrum neutrons to simulate terrestrial exposure and heavy ions to simulate a space environment. More than 10 GBit of cells were irradiated with neutrons and 5 Gbit with heavy ions, providing the ability to examine very low bit error rates. In these studies data errors were measured pre- and postexposure to the dose, with the change in bit error rate (BER) reflecting a disturb in the state of the cell.

While the neutron exposure did not produce a measurable increase in BER (to a fluence of 10^{10} n/cm^2), the heavy ion testing showed a measurable rate of cell disturbs. This was only seen when the angle of incidence was along the wordline direction and was a strong function of the irradiation angle (Fig. 5.31). This was also only seen on cells programmed to the "1" state, which corresponds to the crystalline state of the chalcogenide. No disturbs were measured when the angle of incidence was along the bitline direction.

Further analysis of the wordline disturbs was performed to understand the failure mechanism. The variation of disturb BER with ion exposure at the most critical direction (60° to the wordline) is shown in Fig. 5.32. This indicates that the rate of disturb is only significant at very high LET levels, in this case greater than 40 MeV cm^2/mg. Furthermore, based on the fact that only the SET state was disturbed, the mechanism is likely related to the amorphization of the chalcogenide in the memory cell. (As there was no RESET disturb, there is no indication that the ion exposure is resulting in crystallization.) One possible explanation for this mechanism is the amorphization of a thin layer of chalcogenide at the top interface of the heater element, as depicted in Fig. 5.32. This was reported to have been observed in microscopic images of disturbed cells. Furthermore, modeling of the "track radius" for ions at 60 MeV cm^2/mg suggests that an amorphized region of 12 nm could be created, sufficient to disturb the state of the memory cell. While the LET level required to induce this mechanism does not present a risk of data disturb in terrestrial applications, this analysis indicates that it should be evaluated as the cell is scaled to smaller dimensions where the sensitivity may increase.

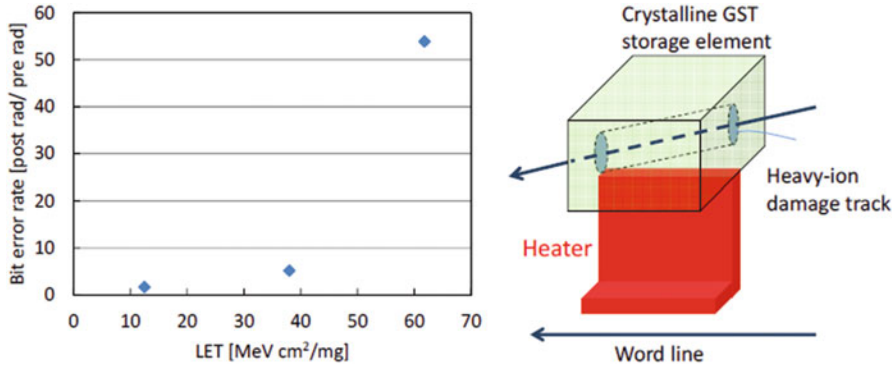


Fig. 5.32 The BER increases with amount of heavy ion (iodine) exposure (the direction of irradiation is at 60° to the wordline direction). The sketch on the *right* shows the direction of the exposure with regard to the memory cell structure [25]

5.8 Error Management Considerations

As the size of memory products continues to increase and cell size decreases, it becomes increasingly difficult to create a "perfect" array of cells. As with other semiconductor devices, many defect mechanisms can be detected during device testing and repaired with redundant elements (e.g., in the case of misshapen features due to particle defects during manufacturing). Other mechanisms can be managed during operation with a sufficiently advanced memory controller. Extreme variance in write cycles across the cells can be reduced by wear leveling, as is commonly done in NAND flash. While these techniques are effective for defects that act predictably, they are not as effective when applied to mechanisms that are erratic in nature. For PCM, such erratic failures can result from the percolative nature of the crystallization and growth processes (as discussed in the prior data retention section). Performing a bake retention test and removing the cells that fail through a repair operation would not be effective in guaranteeing defect-free field operation.

One strategy for addressing the erratic nature of the PCM failures is to implement a process by which defective cells can be fixed "on the fly." Error-correcting code (ECC) schemes provide this capability and are used in a wide range of applications where the defect levels are predictable but the specific "position" of the error is unknown [27]. ECC is widely used in NAND flash-based devices, so the application to semiconductor memory is well established. In the NAND implementations, however, the correction is typically based on a very large code word (greater than 4 kbit) and decoded with a high latency. This scheme is consistent with NAND's intrinsic read latency and provides a very efficient and powerful correction capability. PCM products, however, are generally designed to operate with much lower read latency (~100 ns) and on much smaller data words (typically 128 or 256 bits). As a result, the desired ECC schemes must be optimized for very fast decoding operation. Such a scheme is described by Amato et al., who

BCH ECC key features	
Number of data bits	256
Number of parity bits	18
Number of correctable errors	2
ECC latency overhead	< 10ns
ECC area overhead	< 5% die area

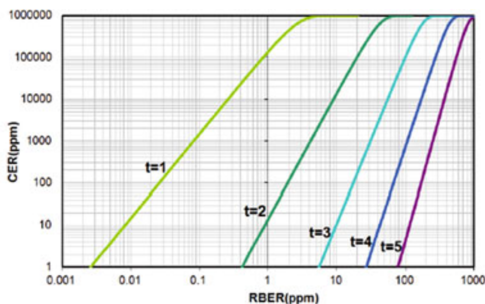


Fig. 5.33 Characteristics of a low-latency, small code word ECC scheme as implemented on a 1Gb PCM component [28]

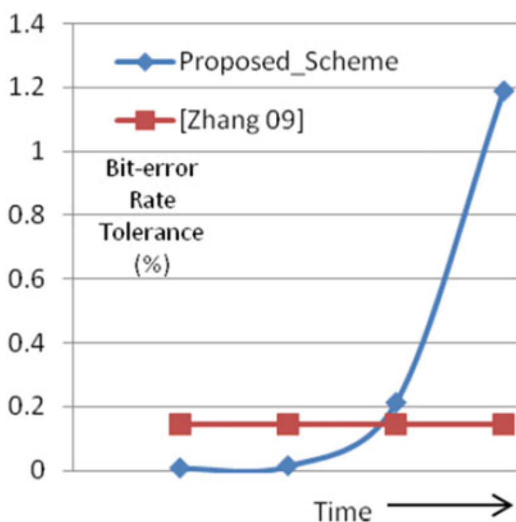


Fig. 5.34 Bit error rate tolerance vs. time for the adaptive ECC scheme as proposed by Datta et al. [29]

used a BCH scheme with 256 data bits and 18 parity bits [28]. This scheme was able to correct 2 errors in the code word in less than 10 ns, as shown in Fig. 5.33a. Figure 5.33b shows the effectiveness of this scheme. Here the y-axis is the chip error rate for a 1 Gb device as a function of the cell raw bit error rate (RBER). Each line corresponds to the level of correctability (cells correctable per code word). For the case of $t = 2$, a product failure rate of less than 10 DPM can be achieved with a cell RBER of up to 10^{-5}. This is 100x higher than the level that could be tolerated with a 1 cell per code word ($t = 1$) scheme and many orders of magnitude higher than what would be acceptable without ECC.

Beyond the fixed-code word ECC scheme described above, adaptive schemes can be used to provide higher levels of error correction (though at a cost of higher complexity). Datta et al. proposed a scheme by which the ECC code word is adjusted to include additional check bits as the base RBER increases [29]. While this leads to an effective reduction in the size of the available memory, it is able to

correct defect levels in the range of 1% (Fig. 5.34). This type of scheme, however, requires the system to monitor the increasing RBER and adjust the code word size. Such a process may be effective for defects introduced when data are written by implementing a verification operation. On the other hand, in cases where the RBER increases without being detected by a read operation (e.g., due to cells failing by data retention or thermal disturb), the system will not be able to adjust the code word, and the method would be less effective.

References

1. B. Gleixner et al., Data Retention Characterization of Phase-Change Memory Arrays, 2007 IRPS
2. A.L. Lacaita et al., Electrothermal and phase-change dynamics in chalcogenide-based memories, 2004 IEDM
3. A. Redaelli et al., Impact of crystallization statistics on data retention for phase change memories, 2005 IEDM
4. D. Mantegazza et al., Statistical analysis and modeling of programming and retention in PCM arrays, 2007 IEDM
5. M. Rizzi et al., Intrinsic retention statistics in phase change memory (PCM) arrays, 2013 IEDM
6. P. Zuliani et al., Overcoming temperature limitations in phase change memories with optimized $Ge_xSb_yTe_z$. Transactions on Electron Device **60**(12), 4020–4025 (2013)
7. G. Navarro et al., Trade-off Between SET and Data Retention Performance Thanks to Innovative Materials for Phase-Change Memory, IEDM 2013
8. JEDEC specifications are available at www.JEDEC.org
9. N. Papandreou et al., Multilevel Phase-Change Memory, ICECS 2010
10. A. Pirovano et al., Reliability study of phase-change nonvolatile memories. IEEE Trans. Device Mater. Reliab. **4**(3), 422–427 (2004)
11. T.Y. Yang et al., Effects of doping on the electric-field-induced atomic migration and void formation in Ge2Sb2Te5, in *18th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2011
12. K. Kim, S.J. Ahn, Reliability investigations for manufacturable high density PRAM, IRPS 2005
13. P.-Y. Du et al., The impact of melting during reset operation on the reliability of phase change memory, in *IEEE International Reliability Physics Symposium (IRPS)*, 2012
14. M.H. Lee et al., The Impact of Hole-induced Electromigration on the Cycling Endurance of Phase Change Memory, 2010 IEDM
15. U.S. Patent 8,036,016, http://www.uspto.gov/
16. W.S. Khwa et al., A Novel Inspection and Annealing Procedure to Rejuvenate Phase Change Memory from Cycling-Induced Degradations for Storage Class Memory Applications, IEDM 2014
17. C. Villa et al., A 45nm 1Gb 1.8V Phase-Change Memory, ISSCC 2010
18. D. Ielmini, Y. Zhang, Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices. J. Appl. Phys. **102**, 054517 (2007)
19. A. Pirovano et al., Scaling analysis of phase-change memory technology, IEDM 2003
20. U. Russo et al., Modeling of programming and read performance in phase-change memories—Part II: Program disturb and mixed-scaling approach. IEEE Transaction on Electron Devices **55**(2) (2008)

21. S. Hudgens, Phase Change Memory: Materials and Device Physics, tutorial presentation at the 2009 International Memory Workshop
22. M. J. Kang et al., PRAM cell technology and characterization in 20nm node size, 2011 IEDM
23. A. Redaelli et al., Interface Engineering for Thermal Disturb Immune Phase Change Memory Technology, 2013 IEDM
24. J. Maimon et al., Chalcogenide memory arrays: Characterization and radiation effects. IEEE Trans. Nucl. Sci. **50**(6), 1878–1884 (2003)
25. S. Gerardin et al., Heavy-Ion Induced Single Event Upsets in Phase-Change Memories. 2014 I. E. International Reliability Physics Symposium
26. G. Servalli et al, A 45nm Generation Phase Change Memory Technology, 2009 IEDM
27. For examples of Error Correcting Codes, see https://en.wikipedia.org/wiki/Error_detection_and_correction
28. P. Amato et al., Ultra fast, Two-Bit ECC for Emerging Memories, 2014 I.E. 6th International Memory Workshop (IMW)
29. R. Datta et al., Designing a Fast and Adaptive Error Correction Scheme for Increasing the Lifetime of Phase Change Memories, 2011 29th IEEE VLSI Test Symposium

Chapter 6
Structure and Properties of Chalcogenide Materials for PCM

Pierre Noé and Françoise Hippert

Abbreviations

AIMD	Ab initio molecular dynamics
AIST	Ag- and In-doped Sb_2Te
ALD	Atomic layer deposition
CD-RW	Compact disc rewritable
CMOS	Complementary metal-oxide-semiconductor
CN	Coordination number
CVD	Chemical vapor deposition
DC	Direct current
DCD	Diffractive coherent domain
DFT	Density functional theory
DoS	Density of states
DRAM	Dynamic random access memory
DSC	Differential scanning calorimetry
DVD	Digital versatile disc
DVD-RAM	Digital versatile disc random access memory
DVD-RW	Digital versatile disc rewritable
EXAFS	Extended X-ray absorption fine structure
fcc	Face-centered cubic
FTIR	Fourier-transform infrared spectroscopy
FWHM	Full width at half maximum

P. Noé (✉)
Univ. Grenoble Alpes, CEA, LETI, MINATEC Campus, Grenoble, France
e-mail: pierre.noe@cea.fr

F. Hippert
LNCMI (CNRS, Univ. Grenoble Alpes, UPS, INSA), Grenoble, France

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_6

GST	Alloys located on the pseudo-binary line connecting GeTe and Sb_2Te_3 in the ternary Ge-Sb-Te phase diagram
MD	Molecular dynamics
MLC	Multilevel cell
MOCVD	Metal-organic chemical vapor deposition
PCM	Phase change memory
PCRAM	Phase change random access memory
PDF	Pair distribution function
PECVD	Plasma-enhanced chemical vapor deposition
PVD	Physical vapor deposition
RAM	Random access memory
RF	Radio frequency
SEM	Scanning electron microscopy
STEM	Scanning transmission electron microscopy
TEM	Transmission electron microscopy
T_g	Glass transition temperature
T_{melt}	Melting temperature
T_x	Crystallization temperature
VDoS	Vibrational density of states
XANES	X-ray absorption near-edge structure
XAS	X-ray absorption spectroscopy

6.1 Introduction

As discussed in previous chapters, some chalcogenide materials can be quickly and reversibly switched between an amorphous and a polycrystalline phase with very different optical and electrical properties (Figs. 6.1 and 6.2). This unconventional property combination led to their use for data storage applications in rewritable optical storage products since the 1990s and in non volatile resistive memories since 2005 (Fig. 6.3). The existence of a reversible transition between a highly resistive state and a conductive state induced by application of an electrical field was first reported in 1968 by Ovshinsky [125] in a chalcogenide material containing Si, Ge, As, and Te. Following this pioneering report, alloys located on the pseudo-binary line connecting GeTe and Sb_2Te_3 in the ternary Ge-Sb-Te phase diagram (GST alloys) were found to present a very fast crystallization and a pronounced change of optical reflectivity between the amorphous phase (low reflectivity state) and the crystalline phase (high reflectivity state) [167, 168]. The prototypical $Ge_2Sb_2Te_5$ alloy, which is the most studied chalcogenide phase change material, belongs to this line. Another family consists of Ag-/In-doped Sb_2Te alloys (AIST alloys), for instance, $Ag_5In_5Sb_{60}Te_{30}$. AIST and GST alloys have been widely used in digital versatile discs (DVD)-RAM, Blu-ray discs, and DVD-rewritable discs [54, 164, 170]. Since the crystallization of GST alloys is accompanied by a decrease

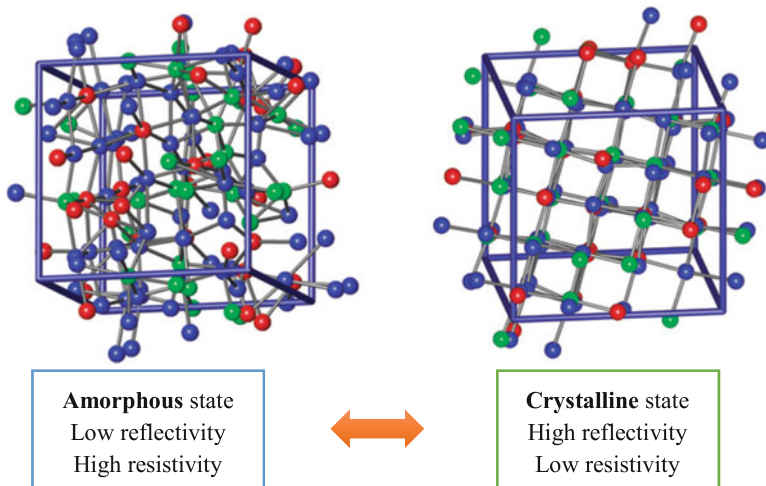


Fig. 6.1 Snapshots of configurations of 63-atom models of amorphous and crystalline $Ge_2Sb_2Te_5$ from [48]. In devices, the amorphous and crystalline states are used to encode information

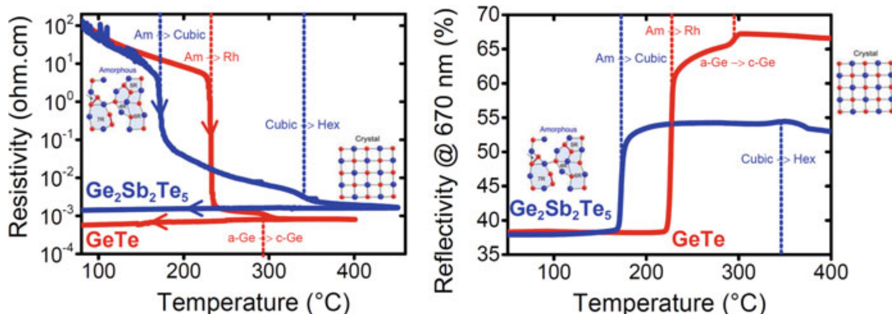


Fig. 6.2 Resistivity (left) and reflectivity (right) measured as a function of increasing temperature (ramp rate of 10 °C/min) for 100 nm thick GeTe (exact composition $Ge_{52.5}Te_{47.5}$) and $Ge_2Sb_2Te_5$ films deposited by DC magnetron sputtering on a thermally oxidized Si substrate and capped in situ with a 10 nm SiN protective layer. The resistivity measured during cooling down is also shown. The amorphous phase (initial state) is characterized by a high electrical resistivity and a low optical reflectivity and the crystalline phase by a much lower resistivity and higher reflectivity. Crystallization (at 173 °C for $Ge_2Sb_2Te_5$ and 234 °C for GeTe) is accompanied by an abrupt change of resistivity and reflectivity. The crystallization temperature and the electrical and optical contrasts are higher for GeTe than for $Ge_2Sb_2Te_5$

of their electrical resistivity by three to four orders of magnitude, these alloys are now also successfully used in phase change memories (PCMs).

The main phase change materials with properties suitable for memory applications are depicted in Fig. 6.3. Most of them are chalcogenide materials containing Te (Fig. 6.4), but non-chalcogenide materials such as Ge-Sb [134] and Ga-Sb [138] alloys also exhibit a phase change behavior. The specificity of phase change

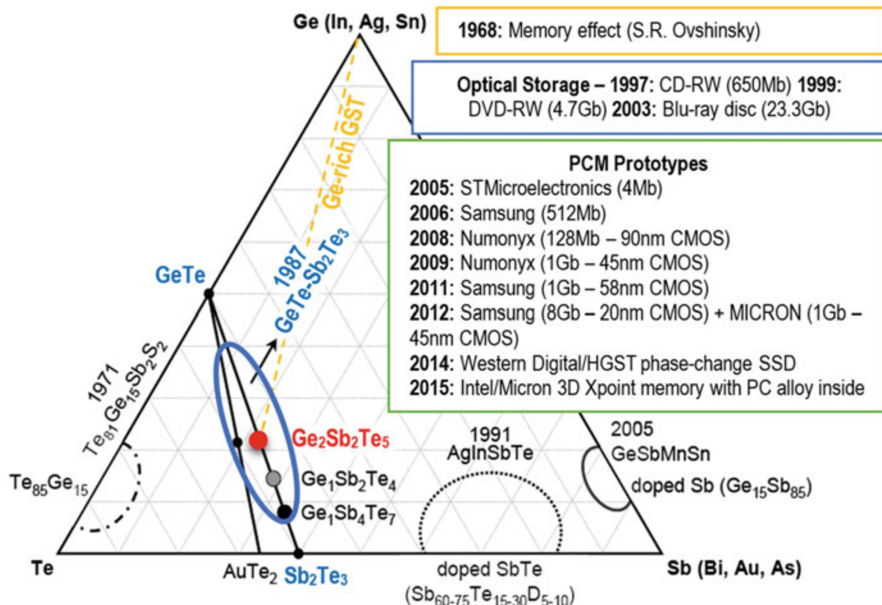


Fig. 6.3 Ternary phase diagram depicting different phase change alloys with their year of discovery and a summary of the development of optical and resistive phase change memories (Adapted from [164]). Many studies have been devoted to alloys located on the pseudo-binary GeTe-Sb_2Te_3 tie line in the ternary Ge-Sb-Te system, including the prototypical $Ge_2Sb_2Te_5$ alloy. More recently, Ge_xTe_{1-x} and $Ge_2Sb_2Te_5$ alloys doped with either B or C, N, O, SiC, SiN, SiO_2, etc. as well as Ge-rich Ge-Sb-Te alloys have been investigated for use in embedded phase change memories

materials is illustrated in Fig. 6.5 where chalcogenides and other compounds are displayed in a map according to their tendency toward s-p hybridization and bond ionicity [91, 92]. Strikingly, phase change materials are located within a very small region of this map. Their crystalline state is characterized by a pronounced resonant character of the bonds (“resonant bonding”) [150]. Suppression of resonant bonding by disorder in the amorphous state could explain the large difference of optical and electrical properties between the amorphous and crystalline phases.

In a PCM cell, the amorphous or crystalline state of a small confined volume of the phase change material is selected by application of a programming current pulse. The procedure is illustrated in Fig. 6.6 for a typical mushroom-type cell. To amorphize the active volume, the temperature must first rise above the melting temperature of the material (T_{melt}) and then decrease fast enough below the glass transition temperature (T_{g}) to impede crystallization. This is achieved by applying a short and intense current pulse (RESET pulse). Crystallization of the active volume can be obtained by heating at a temperature less than T_{melt} but higher than T_{g} to ensure a high atomic mobility and hence a fast crystallization. This is achieved by applying a current pulse (SET pulse) with a longer duration and a smaller intensity

Fig. 6.4 Chalcogenides are materials containing S, Se, or Te (belonging to group 16) alloyed with electropositive elements (often a group 15 (As, Sb) element and/or a group 14 (Ge, Sn) element). Most phase change materials contain Te, usually alloyed with Ge and/or Sb

than the RESET pulse. Note that a debate is still ongoing to decide whether the applied electrical field influences crystallization in a resistive memory cell [65, 93].

Performances of PCM devices are intimately linked to the properties of the integrated phase change material. The main requirements that phase change resistive memory devices must meet to be competitive with flash memories and other emerging non volatile memory technologies are summarized in Table 6.1. Some of these requirements are conflicting. For instance, increasing the crystallization temperature (T_x) to enhance the stability of the amorphous phase has been achieved either by changing the composition of Ge-Sb-Te alloys away from the pseudo-binary line [14, 22, 23, 34, 112, 129, 136, 156, 180] or by doping with elements such as B, C, N, O, SiC, SiN, or SiO_2 [10, 11, 35, 149, 176, 177]. However, the obtained T_x increase is often accompanied by an increase of the crystallization time [149].

A huge research effort has been and is still devoted to phase change materials [15, 28, 77, 78, 92, 135, 164]. Material science is involved to understand the property contrast and the phase transformations (especially the crystallization mechanism) and to optimize materials for specific memory applications, the two issues being intimately linked. Hereafter, we focus mainly on the stability of the amorphous state and on the impact of confinement (interface effects) on

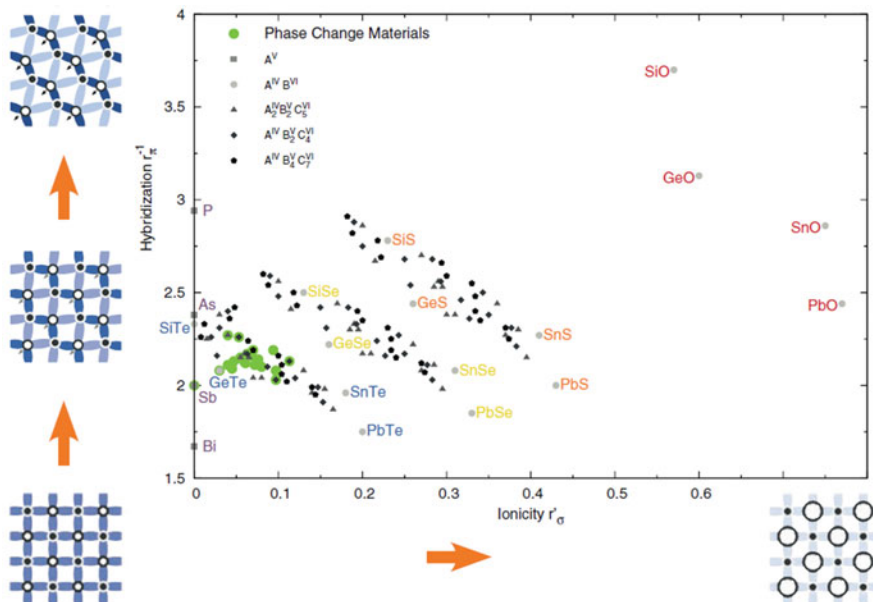


Fig. 6.5 Map from [92] with group V elements (squares), binary $A^{IV}B^{VI}$ compounds (circles), and ternary alloys with different compositions including $A_2^{IV}B_2^{V}C_5^{VI}$ (triangles), $A^{IV}B_2^{V}C_4^{VI}$ (diamonds), and $A^{IV}B_4^{V}C_7^{VI}$ (pentagons). Bands of oxides, sulfides, selenides, and tellurides are clearly discernible and are denoted by red, orange, yellow, and blue letters to label the corresponding binaries. The *x* axis coordinate measures the ionicity of bonds, and the *y* axis coordinate measures the tendency toward s-p hybridization. The schematic structures drawn outside the graph illustrate how the bonding mechanism varies with the coordinates. Phase change materials, marked by green circles, are all found in an area of small ionicity and a limited degree of hybridization, which enables resonance bonding in the crystalline phase

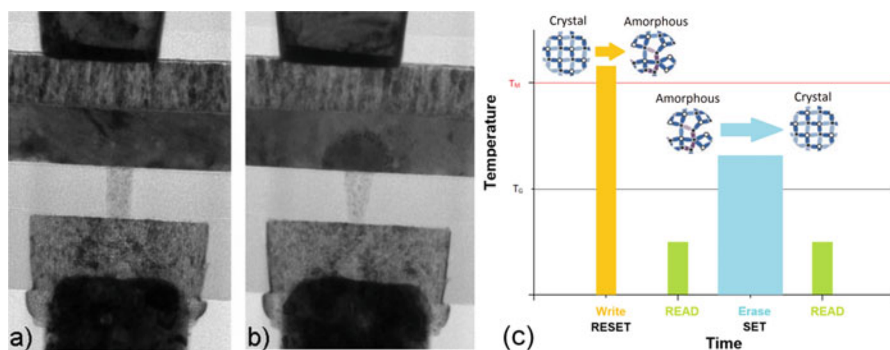


Fig. 6.6 Transmission electron microscopy (TEM) cross sections of a mushroom-type PCM cell in (**a**) the SET state and (**b**) the RESET state where an amorphous dome has been formed inside a crystalline GST layer (Reproduced from [13]). The dome is located above the bottom metallic heater. (**c**) Schematic representation of the electrical programming procedure (RESET and SET pulses). Reading is performed by measuring the resistance of the cell by applying a low intensity current pulse, which does not change the cell state

Table 6.1 Link between device performances and optimization of phase change material properties

Device performance	Phase change material property
High contrast between logical states	Large difference between the electrical resistivity of the amorphous and crystalline states
Low power consumption	Moderate melting temperature coupled with a low thermal conductivity and a not too low resistivity in the crystalline state to reduce the RESET current required for amorphization Control of the threshold switching in the amorphous state
High switching speed	Fast crystallization
High data retention to retain the stored information for a long time (typically 10 years at the operating temperature) and during soldering reflow process during fabrication	Increase the crystallization temperature (T_x) to impede spontaneous crystallization of the amorphous phase at the operating or soldering temperatures Reduce the resistance drift (resistance increase with time, e.g., aging) in the amorphous state
High cyclability (endurance) of devices (at least 10^6 write/read cycles) with low variability of electrical parameters (cell resistance states, programming, and read currents)	Reduce the density difference between the amorphous and crystalline states to reduce the formation of voids and stress buildup in the switched volume Limit element electromigration under high programming electrical fields Limit interdiffusion and/or chemical reaction between phase change material and surrounding materials, such as electrodes and insulating layers Limit element/phase separation at crystallization
High scalability to increase storage density and to allow integration at the smallest CMOS technological nodes	Keep and, if possible, improve the phase change material properties when the volume of the active material decreases down to dimensions at which interface and size effects play a major role

crystallization. These issues are related to the major challenges faced by PCM technology to be competitive with other emerging non volatile memories, beyond its already well-established performances. Phase change material must guarantee code integrity during the technological integration process flow, in particular during the soldering reflow process (2 min at 260 °C) [56]. Besides, embedded PCM technology would constitute a real breakthrough for both process cost savings and performances. For embedded applications, the information must be stored for a long time at elevated temperatures. For example, automotive applications require that the amorphous phase remains stable for 10 years at 150 °C, which excludes the prototypical $Ge_2Sb_2Te_5$ alloy (reported crystallization temperature $T_x \approx 150$ °C). During the last decade, many studies have been conducted by industrial and academic groups to find new phase change materials with higher T_x and hence

enhanced stability of the amorphous phase. In particular, the effect of adding dopants has been explored. Another issue related to the stability of the amorphous state of phase change materials is the so-called resistance drift phenomenon (Chap. 2). At a given temperature, the resistivity of the amorphous phase increases with time. Resistance drift has hindered up to now the development of multilevel storage devices. Besides, the high power consumption of PCM devices, resulting from the high programming currents needed to switch the state of the memory cell [15], is a major obstacle for integration of PCMs at the smallest CMOS technology nodes (<28 nm). Complementary approaches have been developed to reduce the power consumption of PCM cells and particularly the amorphization current (RESET pulse). The first one is based on optimization of material composition to reduce the thermal conductivity of the crystalline state [49]. In parallel, it has been demonstrated that optimizing memory cell design, by increasing the geometrical and thermal confinement of the phase change material, allows to enhance heating efficiency and hence to reduce the programming currents [64, 70] (see Chap. 3). However, incorporation of phase change materials into confined structures opens fundamental issues regarding their scalability limit and the impact of confinement on their phase transformations.

The chapter is organized as follows. Section 6.2 provides an overview of the outstanding physical properties of phase change materials, in relation with their complex atomic structure, by focusing on the two prototypical alloys $Ge_2Sb_2Te_5$ and GeTe. Section 6.3 is devoted to two material issues related to the stability of the amorphous state of phase change materials. We first review how dopants affect the structure of phase change materials by focusing on the case of C- and N-doped GeTe (Sect. 6.3.1). Then, we discuss the structural relaxation of the amorphous phase during aging in connection with the resistance drift issue (Sect. 6.3.2). In Sect. 6.4, we review the state-of-the-art understanding of interface effects on the properties of phase change materials. Scaling effects are treated in Chap. 8. A summary is given in Sect. 6.5, with emphasis on pending material issues that need to be solved for designing and engineering PCM devices with optimized properties and structures.

6.2 Overview of Phase Change Material Properties

6.2.1 Ge-Sb-Te Alloys

Alloys located on the GeTe-Sb_2Te_3 pseudo-binary tie line (Fig. 6.3) have attracted much attention since the pioneering work of Yamada et al. [167, 168]. Stable compounds exist only for well-defined compositions, usually written as $(GeTe)_n(Sb_2Te_3)_m$ (with n and m integers), as shown in Fig. 6.7. They belong to the trigonal crystal system but are often denoted as hexagonal phases in the literature. They are characterized by a periodic stacking of planes perpendicular

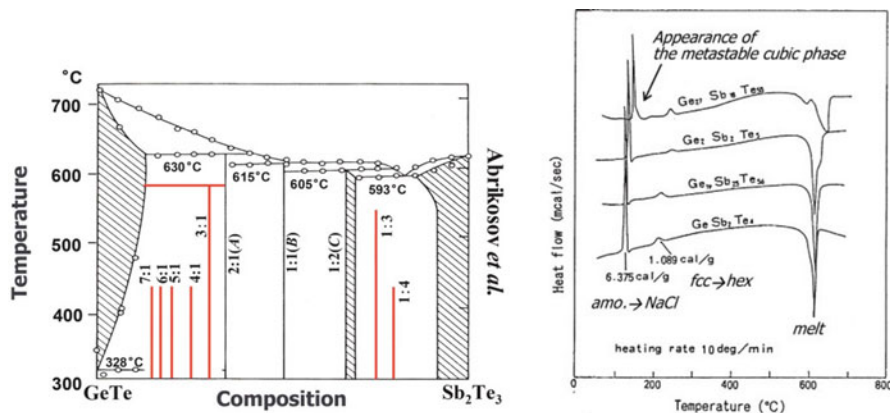


Fig. 6.7 Left: GeTe-Sb_2Te_3 pseudo-binary phase diagram. The liquidus curve is rather flat lying near 600 °C in a wide composition range. Vertical lines indicate the stable $(GeTe)_n(Sb_2Te_3)_m$ compounds, labeled by their n and m values. For instance, 2:1 denotes the $Ge_2Sb_2Te_5$ compound. The first discovered compounds, $Ge_2Sb_2Te_5$ (A), $Ge_1Sb_2Te_4$ (B), and $GeSb_4Te_7$ (C), are indicated by black lines. Other compounds are indicated by red lines. Right: Differential scanning calorimetry (DSC) results, measured on initially amorphous films at a heating rate of 10 °C/min, showing the appearance of the metastable cubic phase (fcc NaCl type), the cubic to hexagonal transition and melting (Reproduced from [170])

to the threefold axis, for instance, 9 planes for $Ge_2Sb_2Te_5$ ($n = 2$, $m = 1$, space group $P\bar{3}m1$) and 21 planes for $GeSb_2Te_4$ ($n = 1$, $m = 1$, space group $R\bar{3}m$) [81, 100, 101]. There are planes fully occupied by Te atoms. The occupation of the other planes by either one single element (Ge or Sb) or a random mixture of Ge and Sb is debated.

However, for ternary alloys on the pseudo-binary tie line, the phase obtained by crystallization of the amorphous phase is either a cubic phase (rocksalt type, space group $Fm\bar{3}m$) (Fig. 6.8) or a rhombohedral phase (space group $R3m$), which can be described as a distorted rocksalt structure. If the alloy composition is expressed as $(GeTe)_{1-x}(Sb_2Te_3)_x$, the cubic phase is obtained for x varying continuously from 0.14 to 0.67 and the rhombohedral phase for lower x values [102, 103]. An outstanding feature of the cubic and rhombohedral ternary phases is the presence of a large number of vacancies [102, 165]. In a rocksalt crystal, Cl and Na occupy the 4(a) and 4 (b) sites, respectively. In the $(GeTe)_{1-x}(Sb_2Te_3)_x$ ternary phases, the 4(a) site is entirely occupied by Te, while the 4(b) site is randomly occupied by Ge, Sb, and vacancies (Fig. 6.8). The vacancy proportion on the 4(b) site is equal to $x/(1 + 2x)$, i.e., 20% in $Ge_2Sb_2Te_5$ [102].Vacancies contribute to lower the energy of the crystalline phases [165]. Their presence also explains why the composition of the crystalline phase (determined by the x value) can vary continuously. This feature enables identical compositions in the crystalline and amorphous phases, which is beneficial for the cyclability of memory devices. The cubic phase is metastable. It is transformed irreversibly into the hexagonal phase by heating (Fig. 6.7). In the case of the binary end compound GeTe ($x = 0$), the rhombohedral phase obtained by

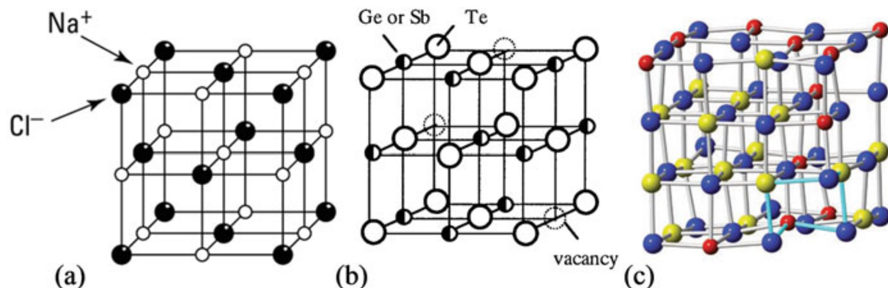


Fig. 6.8 (**a**) Rocksalt cubic unit cell. Cl occupies the 4(a) site and Na the 4(b) site. (**b**) Unit cell of the metastable cubic $Ge_2Sb_2Te_5$ phase (Reproduced from [169]) deduced from X-ray diffraction [101]. The unit cell edge is equal to 0.603 nm at room temperature. The 4(a) site is entirely occupied by Te, while Ge, Sb, and vacancies are randomly distributed on the 4(b) site. The vacancy proportion on the 4(b) site is 20%. In fact, the cubic $Ge_2Sb_2Te_5$ phase is more complex (see Sect. 6.2.2.2) and presents local distortions from the average cubic structure, as illustrated by the model represented in (**c**) (red, Ge; yellow, Sb; blue, Te) from [118]

crystallization of the amorphous phase is stable. To meet the requirements of embedded applications and device fabrication, Ge-rich Ge-Sb-Te alloys with compositions outside the pseudo-binary line have also been studied [23, 112, 156, 180], as illustrated in Fig. 6.3 and Chap. 7, Fig.7.37. However, in the latter case, crystallization is accompanied by a phase separation and segregation of Ge [156].

The atomic structure and the properties of amorphous and crystalline phases of Ge-Sb-Te (GST) alloys belonging to the pseudo-binary line have been the subject of a huge number of studies. An exhaustive review is beyond the scope of this chapter. The interested reader can refer, for instance, to [28, 77, 78, 99] and references therein. In the next sections, emphasis is put on the prototypical $Ge_2Sb_2Te_5$ and GeTe alloys. $Ge_2Sb_2Te_5$ is widely used in resistive memories and is, by far, the most studied phase change material. GeTe is also highly relevant for applications, since it exhibits a fast crystallization rate and a good cyclability [14, 129]. Besides, its binary character simplifies ab initio simulations and interpretation of structural studies. To understand the phase change mechanism and the origin of the property contrast, a key step is to elucidate the atomic structure. Note that phase change materials exhibit somehow paradoxical features. Their fast crystallization suggests minor atomic rearrangements at crystallization. On the other hand, the quite different electrical conductivity and reflectivity of the amorphous and crystalline states suggest significant structural and bonding differences. In the following sections, the structure of the crystalline and amorphous phases of GeTe and $Ge_2Sb_2Te_5$ and their physical properties are reviewed. Then, the crystallization mechanisms are briefly discussed (see also Chaps. 3 and 4).

6.2.2 Structure of $Ge_2Sb_2Te_5$ and GeTe Alloys

6.2.2.1 Methods

To describe the short- and medium-range atomic order in amorphous materials, and more generally in disordered materials, one introduces the total pair correlation function (also called pair distribution function, PDF) $g(r)$, which is the averaged probability of finding one atom at given distance r from another atom [31]. $g(r)$ can be obtained by Fourier transform of the so-called structure factor $S(Q)$, which can be measured by either neutron total scattering or X-ray total scattering (most often by using synchrotron radiation but also in some cases by using high brilliance laboratory equipment). $g(r)$ gives access to interatomic distances (Fig. 6.9) but not to chemical order.

Partial pair distribution functions are introduced to describe the interatomic distance distribution between selected elements. In a binary alloy such as GeTe, there are three partial functions, namely, $g_{Ge\text{-}Te}$ (r), $g_{Ge\text{-}Ge}$ (r), and $g_{Te\text{-}Te}$(r). Measurements of partial PDF are scarce [157] since they require the use of anomalous X-ray scattering. Information on local atomic order around a given element in a compound can be readily obtained by X-ray absorption spectroscopy (XAS) using synchrotron radiation. When the photon energy exceeds the energy of an absorption edge that corresponds to a core electronic level of a given element, the absorption coefficient exhibits oscillations (extended X-ray absorption fine structure, EXAFS), which contain information on the local atomic order around

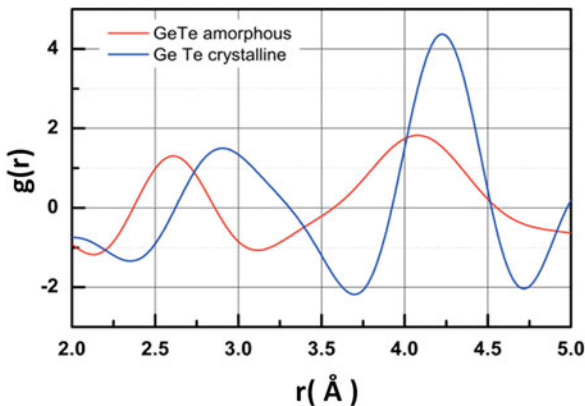


Fig. 6.9 Measured pair distribution function $g(r)$ obtained by Fourier transform of the X-ray total structure factor in amorphous and crystalline GeTe. In crystalline GeTe, the first peak in the range [2.5–3.5 Å] is due to short (2.84 Å) and long (3.16 Å) Ge-Te distances, while the second peak (maximum at 4.23 Å) is due to Ge-Ge and Te-Te distances. A shrinking of all distances is observed in the amorphous phase by comparison with the crystalline phase

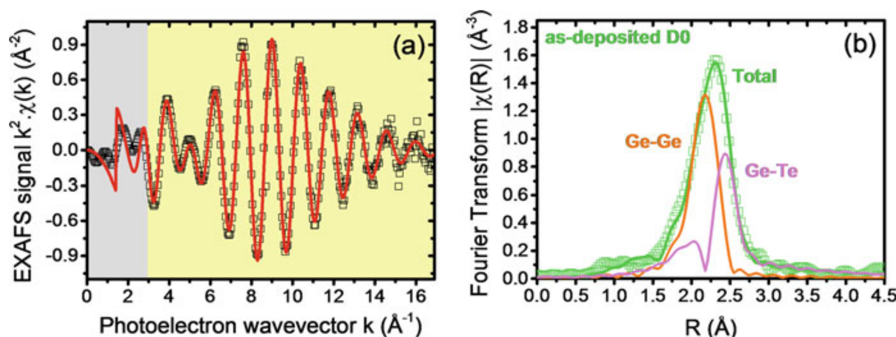


Fig. 6.10 (**a**) k^2-weighted normalized EXAFS oscillations at the K-edge of Ge in an as-deposited amorphous GeTe thin film (squares) and their fit (solid line). *k* is the photoelectron wave vector. (**b**) Fourier transform of the EXAFS oscillations (squares) and their fit (solid line). The partial contributions of Ge and Te backscatters are also drawn. Distances represented in the figure are affected by phase shifts and differ from the real distances, which are determined by the fit procedure (From [115])

the absorbing atom [142] (Fig. 6.10). The EXAFS signal results from interferences of the spherical wave function of the emitted photoelectron with waves backscattered by neighboring atoms. From a fit of the EXAFS signal, one gets the chemical nature and the number of neighbors at a given average distance from the absorbing atom, as well as the distance distribution resulting from thermal (atomic vibrations) and static disorder. For photon energy lower than the absorption edge, the absorption coefficient contains information on the electronic state of the absorbing atom (X-ray absorption near-edge structure, XANES). Total X-ray scattering and X-ray absorption spectroscopy have been widely and successfully applied to investigate the structure of the amorphous state of phase change materials and also to study their crystalline state (see, for instance, [18, 38, 43, 59, 72, 74, 83, 118, 157].) Indeed, the cubic $Ge_2Sb_2Te_5$ phase, as well as other cubic Ge-Sb-Te phases on the pseudo-binary line, presents an appreciable amount of disorder. As a result, their local structure differs from their average structure determined from the analysis of Bragg peaks in standard X-ray diffraction experiments.

Ab initio molecular dynamics (AIMD) methods combining density functional (DF) calculations with molecular dynamics (MD) allow to simulate the amorphous or crystalline states and calculate their electronic properties. The amorphous structure is usually generated by rapid quenching from the melt. By analyzing snapshots of an atomic configuration (Fig. 6.11), one can get information on the local environment of a given atom and calculate partial pair distribution functions (Fig. 6.12) and bond angle distributions. The total structure factor and the total pair distribution function, as well as the EXAFS oscillations, can also be computed and compared to experimental information. Besides, one can obtain information on atom dynamics by calculating either total or partial (for each element or for a specific type of environment of a given element) distributions of vibration frequencies (vibrational density of states, VDoS). This information can be compared to

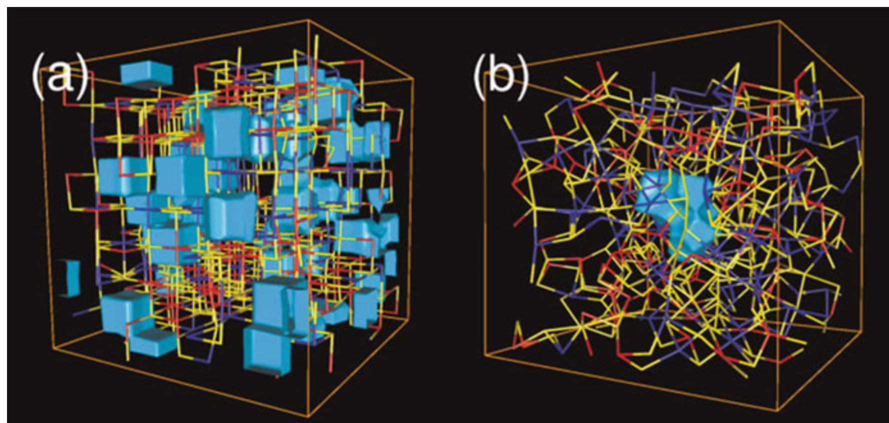
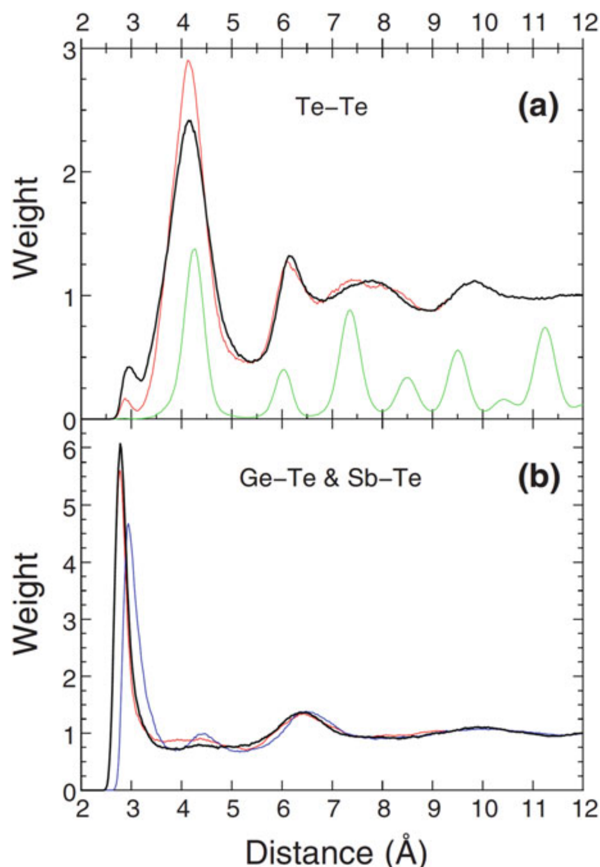


Fig. 6.11 Snapshots of (**a**) crystalline and (**b**) amorphous $Ge_2Sb_2Te_5$ at 300 K (460 atoms and 52 vacancies). Red, Ge; blue, Sb; yellow, Te. Cavities are shown as light blue isosurfaces (From [4])

Fig. 6.12 Simulated partial pair distribution functions (PDF) at 300 K from [1]. (**a**) Te-Te partial PDF in amorphous $Ge_2Sb_2Te_5$ (black), amorphous GeTe (red), and cubic crystalline $Ge_2Sb_2Te_5$ (green, scaled by 0.25). (**b**) Ge-Te (black) and Sb-Te (blue) partial PDF in amorphous $Ge_2Sb_2Te_5$, Ge-Te partial PDF in amorphous GeTe (red)



infrared or Raman spectroscopy data, as illustrated in Fig. 6.13. AIMD simulations led to a detailed insight in the structure and chemical bonding in the amorphous state of phase change materials (for a review, see [4, 99]). Thanks to recent increase in computing capacity and to method improvements, it is now possible to simulate hundreds of atoms over simulation times of the order of a few ns, close to the time involved in crystallization of phase change materials, and hence to investigate crystallization mechanisms (see, e.g., [63, 90, 144, 154, 155] and Sect. 6.2.4).

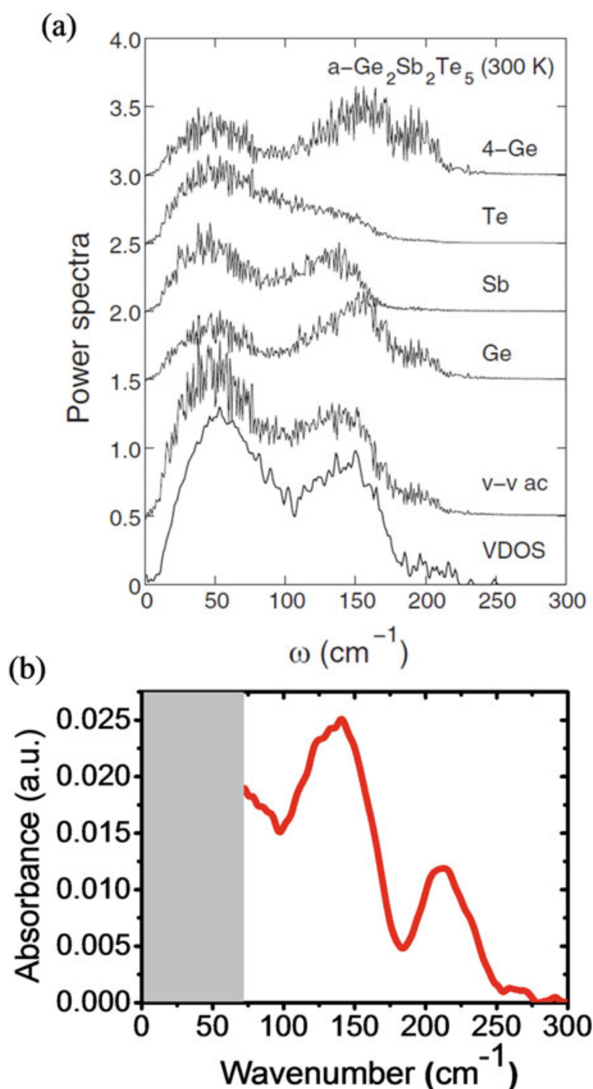


Fig. 6.13 (**a**) Total vibrational density of states (VDoS) for amorphous $Ge_2Sb_2Te_5$ calculated using too distinct methods (from the dynamical matrix for the lowest curve and from velocity autocorrelation function for the curve labeled v-v ac) and partial VDoS (calculated from velocity autocorrelation function) for Ge, Sb, Te, and also Ge in tetrahedral units (curve labeled 4-Ge) from [4]. (**b**) Absorption spectrum measured by Fourier-transform infrared (FTIR) spectroscopy on a 150 nm thick amorphous $Ge_2Sb_2Te_5$ film deposited by DC magnetron sputtering on Si. The simulations well account for the experimental data, given that no direct comparison of amplitudes can be made. The simulation yields the VDoS, while the FTIR absorbance depends on the nature of the vibration modes

6.2.2.2 Crystalline Phases

Amorphous GeTe crystallizes into a rhombohedral phase (Fig. 6.14) with the space group *R3m* [20, 21]. This phase is stable. It has been the subject of numerous studies, regardless of the development of phase change materials, because of its ferroelectric character and the existence of a reversible structural transformation (around 330 °C) to a high-temperature paraelectric cubic phase with the space group $Fm\bar{3}m$. Recent experiments brought a new light on this structural transition [20, 38].

The rhombohedral GeTe phase can be described as a distorted rocksalt structure that has undergone a shear along the cube diagonal ([111] direction), with relative displacements of the Ge and Te atoms along that direction (Fig. 6.15). In this description, the unit cell is a deformed cube with edge equal to 0.5985 nm and angle equal to 88.17° at 295 K [21]. If Te is at the origin of the unit cell, Ge is located off-centered in 0.525, 0.525, and 0.525. The Ge displacement is responsible for the ferroelectric properties of rhombohedral GeTe. Ge is surrounded by six Te atoms forming a distorted octahedron. Due to the Ge off-centering, the Ge-Te distances are split into three short (0.284 nm) and three long (0.316 nm) distances, as illustrated in Fig. 6.15. This splitting is interpreted as resulting from a Peierls distortion as discussed later in Sect. 6.2.3.2. The Ge coordination is denoted as (3 + 3). Note that long Ge-Te bonds alternate with short Ge-Te bonds along [100] ([010] and [001]) cubic directions. The angle between short and long bonds is equal to 171.8°. The structure can also be described as a stacking of Ge and Te planes

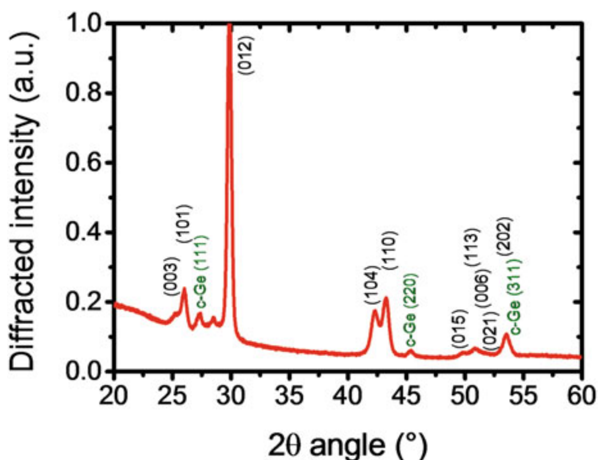


Fig. 6.14 Room temperature X-ray diffraction ($\lambda = 0.1542$ nm) pattern of a 100 nm thick GeTe crystallized film. The film was deposited by sputtering and capped in situ by a 10 nm SiO_2 layer. The as-deposited amorphous film (exact composition $Ge_{52.5}Te_{47.5}$) has then been heated to 400 °C. In addition to the diffraction peaks of the rhombohedral GeTe phase (black labels, hexagonal indexation), weak diffraction peaks originating from a small amount of crystallized fcc germanium (green labels) are also observed

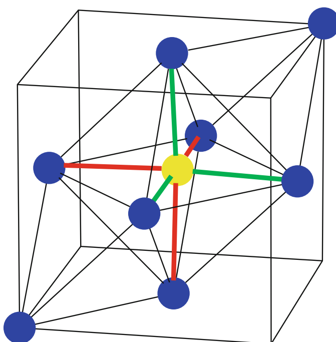


Fig. 6.15 GeTe rhombohedral unit cell (edge equal to 0.4299 nm and angle to 57.93° at 295 K [21]), inscribed into a distorted cube (cube edge 0.5985 nm and cube angle 88.17°). Te atoms (in blue) are located at the corners of the rhombohedral unit cell, and the Ge atom (in yellow) is located off-centered in (0.525, 0.525, 0.525). The first neighbors of Ge are six Te atoms, which form a distorted octahedron. Due to Ge off-centering, there are three short (0.284 nm, in green) and three long (0.316 nm, in red) Ge-Te distances

perpendicular to the [111] direction. One Ge (resp. Te) plane is surrounded by two Te (resp. two Ge) planes. In contrast with the case of a non-distorted rocksalt structure, these planes are not equidistant.

According to the description presented above, the composition of rhombohedral GeTe is exactly $Ge_{0.5}Te_{0.5}$, and there are no vacancies, each crystallographic site being entirely occupied either by Te or by Ge. However, the $Ge_{0.5}Te_{0.5}$ composition lies slightly outside the single-phase region of the rhombohedral phase, which is located on the Te-rich side of the binary GeTe phase diagram [147]. Therefore, there are vacancies on the Ge site in the rhombohedral phase. From Rietveld analysis of powder X-ray diffraction data in a crystallized film [103], the composition of the rhombohedral phase is $Ge_{0.49}\square_{0.01}Te_{0.5}$, where $\square$ denotes vacancies. Nevertheless, in the following, we will continue to label the rhombohedral phase as GeTe. Furthermore, the composition of amorphous films made by sputtering is usually slightly enriched in Ge. Then, one expects rejection of excess Ge atoms outside the rhombohedral phase in crystallized films. This is confirmed by the detection of Bragg peaks ascribed to cubic germanium in X-ray diffraction patterns measured on films heated above ≈ 300 °C (Fig. 6.14). In films heated above the crystallization temperature of GeTe but below 300 °C, excess Ge atoms are expected to form an amorphous germanium phase. This conclusion is supported by the detection of Ge-Ge bonds equal to 0.246 nm in EXAFS experiments in a film annealed at 250 °C [73]. If the overall composition of the film is $Ge_{0.52}Te_{0.48}$ and the composition of the rhombohedral phase $Ge_{0.49}\square_{0.01}Te_{0.5}$, 10% of the Ge atoms would be in a germanium amorphous phase. This estimation is compatible with the proportion of Ge-Ge bonds deduced from EXAFS analysis in [73], but the vacancy concentration in rhombohedral GeTe is not equal to 10% as concluded in [73] but to 1%.

Analysis of Bragg peak positions as a function of temperature allows to detect the reversible structural phase transition between the rhombohedral and cubic phases. Reported transition temperatures vary between 277 °C [103] and 327 or 427 °C [20]. It should be emphasized that only the average structure of the high-temperature phase is cubic. Local structure studies (EXAFS experiments and determinations of the pair distribution function) have established that the local order is unchanged at the structural transition and that short and long Ge-Te bonds continue to exist above the phase transition temperature [38] (Fig. 6.16). This result means that Ge atoms are still off-centered in the high-temperature phase. The average Ge displacement is zero since the directions of the displacements (along any of the cube diagonals) differ in different unit cells.

Amorphous $Ge_2Sb_2Te_5$ crystallizes into a cubic (rocksalt type) metastable phase. The cubic lattice parameter a is equal to 0.6029 nm at room temperature [101]. When heated above ≈ 330 °C, the cubic phase transforms irreversibly into the stable hexagonal phase. In the cubic phase model (Fig. 6.8b) deduced from Rietveld analysis of powder X-ray diffraction pattern [101], the environment of each atom is octahedral, and Ge (or Sb) are surrounded by Te, while Te are surrounded by Ge, Sb, and vacancies. The structure can be described as a stacking of equidistant planes perpendicular to the [111] direction. Te planes are separated by planes containing Ge, Sb, and vacancies. In contrast with the hexagonal stable phase, there are no adjacent Te planes in the cubic phase. However, the cubic structure represented in Fig. 6.8b is an average structure that ignores the existence of local distortions. The latter are detected from EXAFS experiments at the Ge, Sb, and Te K-edges [74]. No Ge-Sb bonds or Te-Te bonds are detected in EXAFS, in agreement with the chemical order deduced from X-ray diffraction. The first Te-Te distance measured by EXAFS is equal to 0.426 ± 0.001 nm, i.e., close to the distance calculated for a face-centered cubic structure (a/√2). However, the Ge-Te (0.283 ± 0.001 nm) and Sb-Te (0.291 ± 0.001 nm) first distances measured by EXAFS are much shorter than $a/2$, which show that Ge and Sb atoms are displaced

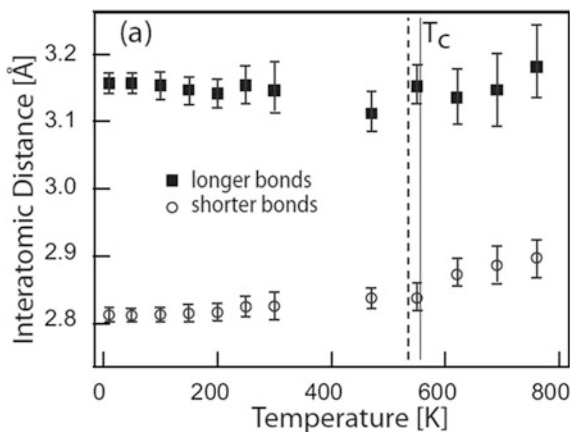


Fig. 6.16 Long and short Ge-Te distances determined by analysis of EXAFS oscillations at the Ge K-edge in crystalline GeTe as a function of temperature (Reprinted from [38]). In the studied film, the rhombohedral-cubic transition occurs at $T_c = 550$ K

away from the ideal rocksalt position (0.5, 0.5, 0.5). Information on the longer Ge-Te and Sb-Te bonds (estimated around 0.32 ± 0.03 nm, i.e., longer than *a*/2, [76]) is less accurate, probably because of broad distance distributions. These results show that the local structure of the cubic $Ge_2Sb_2Te_5$ phase is similar to that of the high-temperature cubic GeTe phase. Ge and Sb are displaced but their displacements are uncorrelated at a long distance. Similar conclusions are drawn from EXAFS studies of the $Ge_1Sb_2Te_4$ alloy [32].

Until recently, experimental results led to the conclusion that vacancies are randomly distributed in cubic $Ge_2Sb_2Te_5$ and in other cubic Ge-Sb-Te phases on the pseudo-binary line although ab initio simulations had suggested that ordering of vacancies could lower the energy [26, 158]. A recent study [12] has brought a new insight. Vacancy ordering in vacancy layers (perpendicular to [111] cubic direction) has been detected in an epitaxial single crystal grown by molecular beam epitaxy (MBE) and in films previously deposited in the amorphous state on a Si (111) substrate and subsequently crystallized by either thermal annealing or application of laser pulses [12]. The degree of vacancy ordering in the crystallized film was found to depend on thermal treatment. Ordering of vacancies has been correlated with resistivity changes in the crystalline $Ge_2Sb_2Te_5$ phase (Sect. 6.2.3.2).

In conclusion, the cubic phase of Ge-Sb-Te alloys on the pseudo-binary line is characterized by a large amount of disorder at the local scale. Such a disorder is not present in rhombohedral GeTe, but the two phases share important common characteristics, i.e., the absence of homopolar Ge-Ge (Sb-Sb) or Te-Te bonds and the existence of short (≈ 0.28 nm) and long (≈ 0.32 nm) Ge-Te bonds resulting in a distorted octahedral (3 + 3) coordination around Ge atoms. To understand the outstanding phase change properties of these materials, these characteristic structural features must be compared with the local structure of the amorphous state.

6.2.2.3 Amorphous Phases

The local atomic order in amorphous $Ge_2Sb_2Te_5$, in other Ge-Sb-Te alloys on the pseudo-binary line, and in binary GeTe amorphous phase is rather similar. As a first step, it is instructive to compare the pair distribution functions $g(r)$ measured by X-ray total scattering in amorphous and crystallized GeTe (Fig. 6.9). In the distance range [0.2–0.5 nm], the peaks of $g(r)$ in the amorphous state occur at smaller distances than in the crystalline state. The effect is particularly strong for the first peak. Let us recall that in crystalline GeTe, this peak is due to short (0.284 nm) and long (0.316 nm) Ge-Te bonds in distorted octahedra. In amorphous GeTe, EXAFS experiments at the Ge K-edge [32, 98, 115] reveal that homopolar Ge-Ge bonds (≈ 0.247 nm) are present in the amorphous phase in addition to Ge-Te bonds (≈ 0.262 nm). Hence, Ge-Ge bonds contribute significantly to the first peak of the pair distribution function $g(r)$ measured by X-ray scattering. All these features are found by ab initio simulations [1, 43], as illustrated in Fig. 6.17. There are no short Te-Te bonds. Similar trends are observed in amorphous $Ge_2Sb_2Te_5$ [1, 3, 4, 17, 72, 74] (Figs. 6.12 and 6.18),

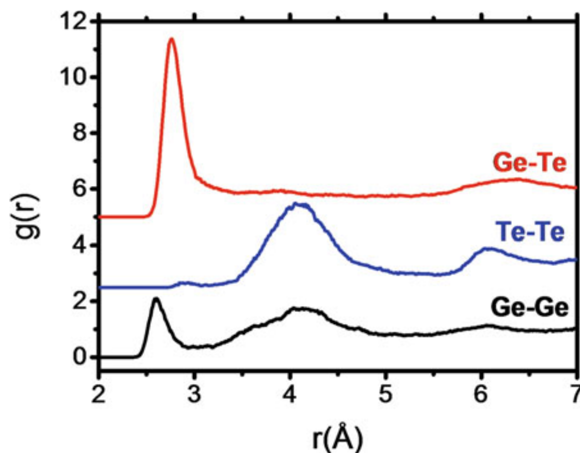


Fig. 6.17 Partial pair distribution functions in amorphous $Ge_{52}Te_{48}$ calculated by ab initio molecular dynamics simulations (Courtesy of J.Y. Raty), using the method described in [139]. In the range [2.4–3 Å], homopolar Ge-Ge bonds are present in addition to Ge-Te bonds, but there are no Te-Te bonds. The calculated distances are slightly larger than observed. This systematic difference arises from the generalized gradient approximation exchange correlation used in the calculations

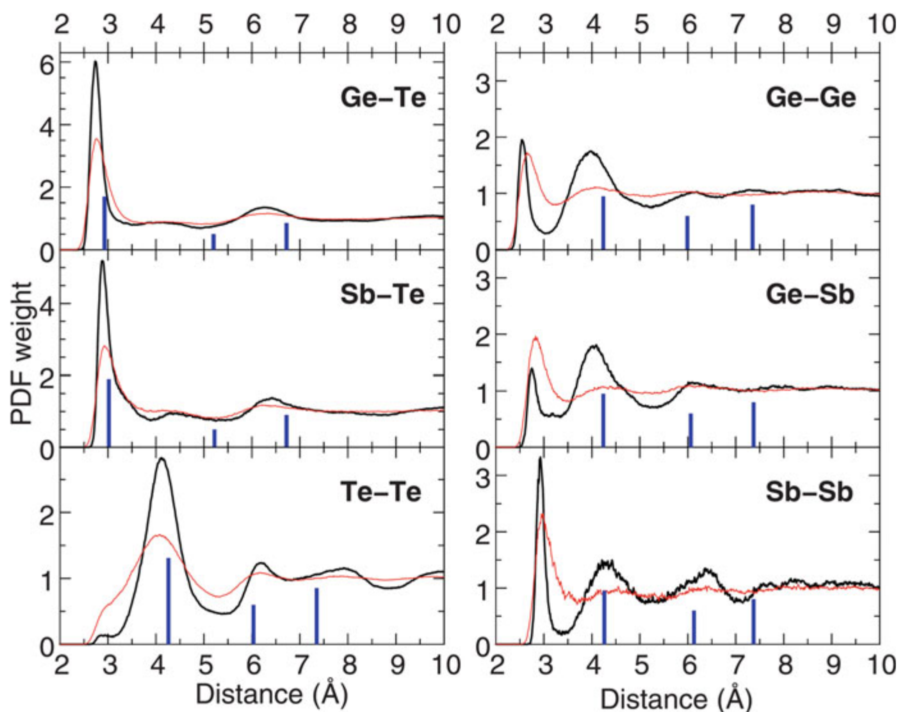


Fig. 6.18 Calculated partial pair distribution functions in $Ge_8Sb_2Te_{11}$ in the amorphous state at 300 K (black lines) and in the liquid at 950 K (red lines) (From [2]). The vertical bars correspond to distances in the crystalline state

$Ge_1Sb_2Te_4$ [139], or $Ge_8Sb_2Te_{11}$ [2]. In the ternary alloys, Sb-Te short bonds also exist, in addition to Ge-Ge and Ge-Te short bonds.

Coordination numbers (CN) around a given element can be obtained by integrating the calculated partial pair distribution functions, but their values depend critically on the integration range, and these numbers should be considered with care. Nevertheless, CN in the amorphous state of GeTe and Ge-Sb-Te are systematically lower than in crystalline state. A remarkable feature of the simulations is the presence of *ABAB* rings (*A*, Ge or Sb, and *B*, Te) in the amorphous state [1, 4, 72], as shown in Fig. 6.19. From analysis of bond angles in simulations, another striking feature emerges. Ge atoms are found to be either in a tetrahedral environment (corresponding to sp^3 bonding) or in a distorted octahedral environment (p bonding) with bond angles close to 90° [2, 17, 139] (Fig. 6.20). It must be emphasized that only distorted octahedral environment exists in the crystalline phases (see above Sect. 6.2.2.2 and Fig. 6.15). Octahedral environments are more distorted in the amorphous phase than in the crystalline phases. Ge, and to a less extent Sb, displacements are larger in the amorphous state. As a result, the short Ge-Te (Sb-Te) bonds shrink, and the long bonds expand in the amorphous phase by comparison with the crystalline phase. Besides, in the amorphous phase, the distribution of long Ge-Te (Sb-Te) bonds in distorted octahedra is so large that

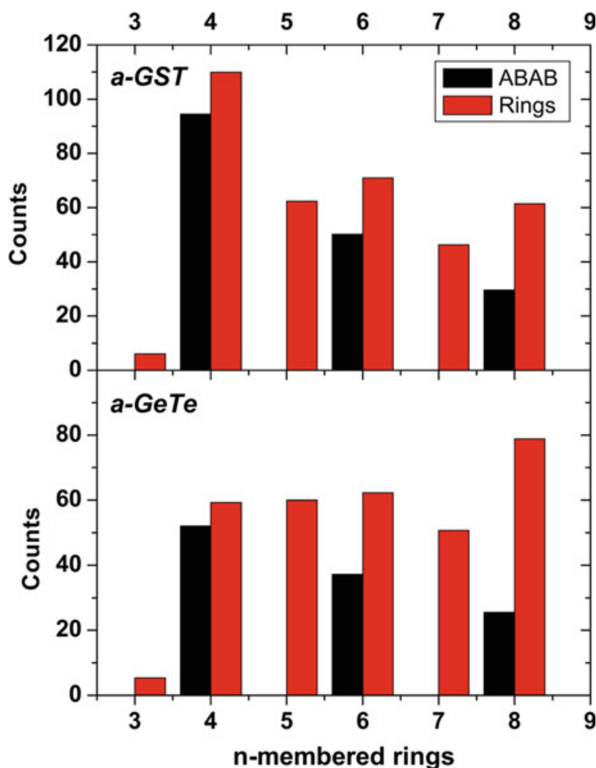


Fig. 6.19 n-fold ring statistics calculated from ab initio simulations of amorphous $Ge_2Sb_2Te_5$ and GeTe at 300 K. *ABAB* configurations refer to alternation of atoms of type *A* (Ge, Sb) and *B* (Te). It is striking to note that most of the four-membered rings have *ABAB* alternation (From [1])

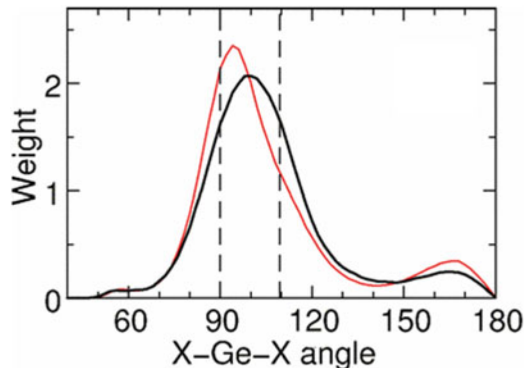


Fig. 6.20 Bond angle distribution around Ge (with bond cut-off equal to 0.32 nm) obtained from ab initio simulations for two different amorphous $Ge_2Sb_2Te_5$ models (From [3]). The red line corresponds to a melt-quenched amorphous state and the black line to an as-deposited amorphous state. Vertical dashed lines denote octahedral (90°) and tetrahedral (109.47°) bond angle values. Tetrahedral environments are present in both models, but there are more tetrahedral Ge atoms in the as-deposited state (58%) than in the melt-quenched one (36%)

these distances are usually not detected in EXAFS experiments. Hence, the Ge-Te or Sb-Te distances deduced from EXAFS experiments correspond to short bonds in octahedral environments. In simulations of melt-quenched amorphous states, about one third of Ge atoms are found in a tetrahedral environment [17, 141], and Ge-Ge bonds in the amorphous state are mostly present in tetrahedral environments. One explanation is provided in [27, 141] where homopolar Ge-Ge bonds are shown to contribute to the local stability of tetrahedra in amorphous GeTe. Experimental proofs of the existence of tetrahedra are brought by Raman spectroscopy [104] and XANES [84].

Whether the structure, and properties, of the amorphous state of phase change materials differs in the as-deposited state (in thin films made by sputtering) and in the melt-quenched state (in devices or in laser-quenched marks in thin films) is a long-standing debate. Indeed, crystallization times for as-deposited materials have been found longer than for melt-quenched materials [137]. Differences in the EXAFS spectra at the Ge K-edge of melt-quenched and as-deposited amorphous $Ge_2Sb_2Te_5$ could be observed, as shown in Fig. 6.21 [18]. The number of Ge-Ge bonds is higher in the as-deposited sample (50 nm thick thin film) than in a melt-quenched sample (obtained by laser irradiation of the previous film). This observation is in agreement with the ab initio simulations reported in [3]. Usually, ab initio simulations start from a very high-temperature ($\approx$ 3000 K) liquid that is cooled to room temperature at a very high cooling rate. Such a state should be close to the melt-quenched amorphous phase in devices. However, a simulation of a sputtered amorphous state could be performed in [3] and compared to a rapidly cooled amorphous model. A striking result is that there are more tetrahedrally coordinated Ge atoms and Ge-Ge bonds in the as-deposited amorphous state than in the rapidly cooled one (Fig. 6.20). These differences could explain the change in

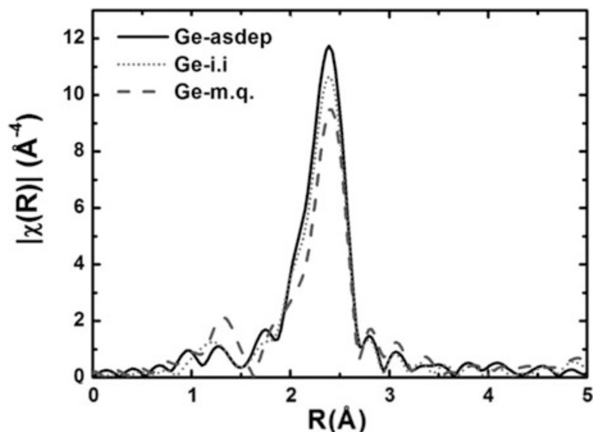


Fig. 6.21 Fourier transform of EXAFS oscillations at Ge K-edge in as-deposited, ion-implanted (i-i), and melt-quenched (m.q.) amorphous $Ge_2Sb_2Te_5$ from [18]

crystallization speeds between as-deposited and melt-quenched samples since homopolar bonds, absent in the crystalline phase, could act as an obstacle to crystallization.

6.2.3 Physical Properties of $Ge_2Sb_2Te_5$ and GeTe Alloys

6.2.3.1 Volume Change Upon Crystallization

Ge-Sb-Te alloys exhibit a large volume reduction (between 6% and 10%, [114, 117]) upon crystallization (Fig. 6.22), which leads to a considerable mechanical stress building in PCM cells and to formation of voids [29] and potentially limits the cyclability and reliability of phase change memory devices. Significant efforts are currently made to identify phase change materials with a smaller volume change upon crystallization. Addition of light doping elements, such as N, C, or O in phase change alloys, primarily done to increase the crystallization temperature (see Chap. 7), has been shown to reduce volume shrinking [172, 176, 178]. For instance, addition of about 5 at % of N in GeTe films suppresses the film thickness change upon crystallization, while keeping a strong electrical contrast between the amorphous and crystalline phases (Fig. 6.23) [172]. Note that similar trends have been observed in a non-chalcogenide phase change material, Ge-doped GaSb [132], but this alloy leads to element segregation upon crystallization.

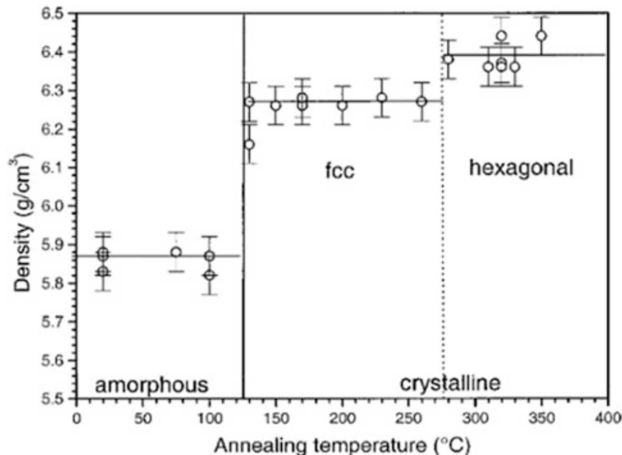
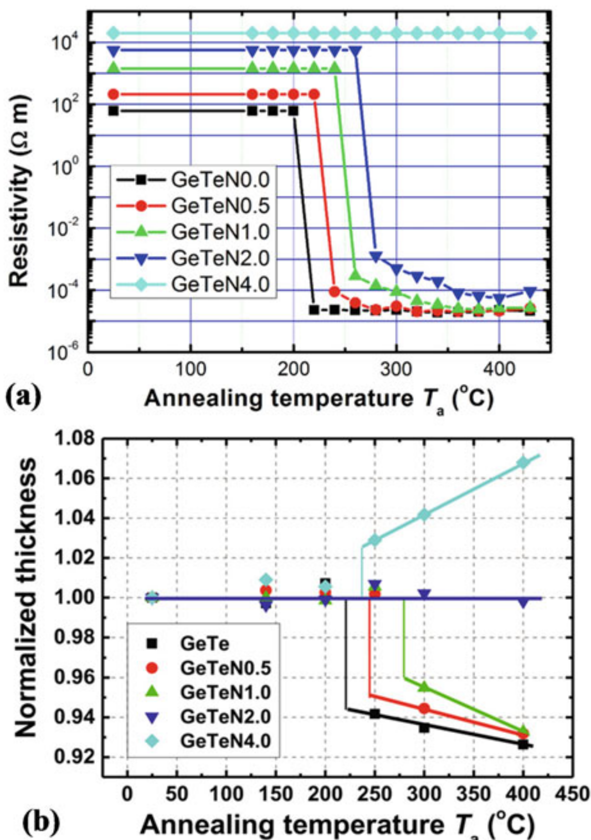


Fig. 6.22 Density of a $Ge_2Sb_{2.04}Te_{4.74}$ film obtained from X-ray reflectivity measurements at room temperature after annealing at different temperatures (From [114]). The relative density increase between the amorphous and cubic fcc phases is equal to 7%

Fig. 6.23 (**a**) Resistivity and (**b**) normalized film thickness at 300 K for GeTe and N-doped GeTe films as a function of the annealing temperature (From [172]). Sample labels refer to the N_2 flux rate during the film deposition by sputtering under N_2/Ar atmosphere. For instance, the sample labeled GeTeN2.0 contains 5 at. % of N



6.2.3.2 Transport and Optical Properties

Resistivity (ρ) measurements of initially amorphous $Ge_2Sb_2Te_5$ and GeTe films measured as a function of temperature are shown in Fig. 6.2. Both materials are semiconductors in the amorphous state [8, 51, 66], and thus the resistivity in the amorphous state smoothly decreases with increasing temperature. Crystallization of $Ge_2Sb_2Te_5$ into the cubic phase and of GeTe into the rhombohedral phase leads to an abrupt resistivity drop (see also, for instance, [39, 46, 151]). A difference between the electrical properties of crystallized GeTe and $Ge_2Sb_2Te_5$ can be observed in Fig. 6.2. Upon heating, after crystallization, the resistivity of GeTe is almost temperature independent. The small accident observed around 300 °C is correlated with the crystallization of the small excess of Ge (see above Sect. 6.2.2.2). Then, when heating up to 400 °C and during subsequent cooling, the resistivity exhibits a metallic behavior ($d\rho/dT$ positive). Detailed studies of the transport properties [8]) establish that rhombohedral GeTe is a degenerate semiconductor, with a large hole concentration and hence a metal-like, *p*-type, conduction. However, band structure calculations predict that a rhombohedral $Ge_{0.5}Te_{0.5}$ crystal would be a narrow gap semiconductor. An explanation of this apparent contradiction is provided in [30], where it is shown that Ge vacancies are responsible for the p-type conduction of rhombohedral GeTe. In $Ge_2Sb_2Te_5$, in contrast to GeTe, the resistivity measured after crystallization in the cubic phase decreases with increasing temperature. A similar behavior is observed in $Ge_1Sb_2Te_4$ and in other cubic Ge-Sb-Te phases on the pseudo-binary line [151] (Fig. 6.24). This strong decrease of the resistivity with increasing temperature has been ascribed to the impact on transport properties of the progressive formation of ordered vacancy layers with increasing temperature [12, 175]. At higher temperature, the metastable cubic phase transforms into the stable hexagonal phase, which is retained when cooling down at room temperature and has a metallic behavior.

A striking feature of transport in crystalline Ge-Sb-Te alloys can be observed in Fig. 6.24. The room temperature resistance of the crystalline phase strongly depends on the maximum temperature reached during the heating process. Hence, decreasing disorder by increasing the annealing temperature allows to decrease the room temperature resistance. This point must be considered in memory cells using Ge-Sb-Te alloys since the resistance contrast is expected to depend on the temperature reached during the SET pulse. This property might be beneficial in future multilevel storage devices.

Besides their outstanding electrical properties, phase change materials also exhibit an unconventional optical contrast between their amorphous and crystalline phases (see, for instance, Fig. 6.2 and [88]), which has led to their use for optical data storage applications. Changes in optical properties between amorphous and crystalline thin films, or as a function of annealing temperature, have been widely studied in literature [7, 87, 111, 120, 121, 126, 153, 162, 166]. Optical constants (refractive index n and extinction coefficient k) and absorption coefficient α as a function of the photon energy are shown in Fig. 6.25 for the amorphous and

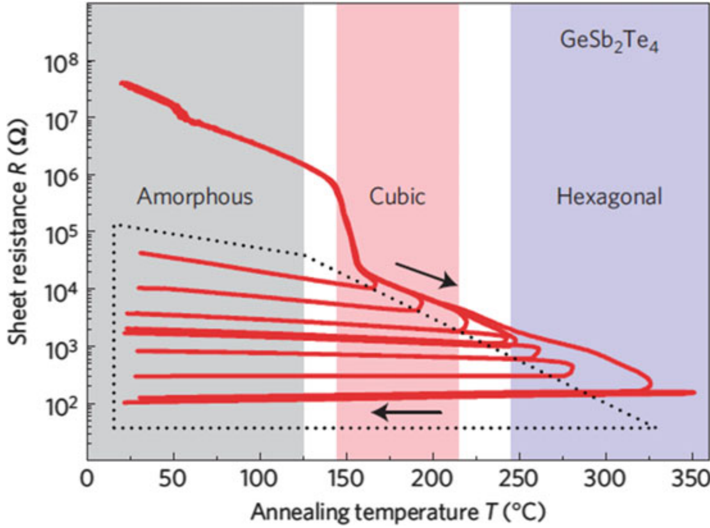


Fig. 6.24 Temperature dependence of the sheet resistance R of a 100 nm thick $GeSb_2Te_4$ film on heating and subsequent cooling to room temperature (From [151]). The abrupt change at 150 °C is due to the crystallization of the initially amorphous film. The room temperature resistance of the crystalline phase strongly depends on the maximum temperature T_{max} reached before cooling. It decreases by two orders of magnitude when T_{max} increases from 170 to 350 °C. Note the change of dR/dT in the crystalline state from a negative to a positive value when T_{max} increases

crystalline rhombohedral phases of GeTe and amorphous and crystalline hexagonal phases of $Ge_2Sb_2Te_5$. The optical constants for the crystalline fcc (not shown here) and hexagonal phases of $Ge_2Sb_2Te_5$ are very similar. A significant change of optical constants between the amorphous and crystalline phase is evidenced. Note that the presence of free carriers in the crystalline phases requires to introduce a Drude contribution to account for infrared metal-like absorption when fitting optical data. From the analysis of the absorption coefficient α as a function of the photon energy, using the $E_g{}^{04}$ method, where $E_g{}^{04}$ is the energy at which the absorption coefficient equals 10^4 cm^{-1}, one can deduce an optical bandgap $E_g{}^{04}$ ~ 0.83 eV for amorphous GeTe and $E_g{}^{04}$ ~ 0.74 eV for rhombohedral GeTe. For $Ge_2Sb_2Te_5$, an optical bandgap (Tauc gap) $E_g{}^{opt}$ of ~0.7 eV for the amorphous phase and ~0.5 eV for the hexagonal crystalline phase has been reported in literature [87, 120, 126]. It must be emphasized that the optical properties of phase change materials highly depend on deposition parameters and film micro-structure [68, 95, 106, 168]. This explains the large discrepancy of optical bandgap values reported in literature.

The origin of the outstanding contrast of optical and electrical properties between the amorphous and crystalline phases of phase change materials has been the subject of many works and is still under debate. From the present understanding of the structure of the amorphous and crystalline states (Sects.

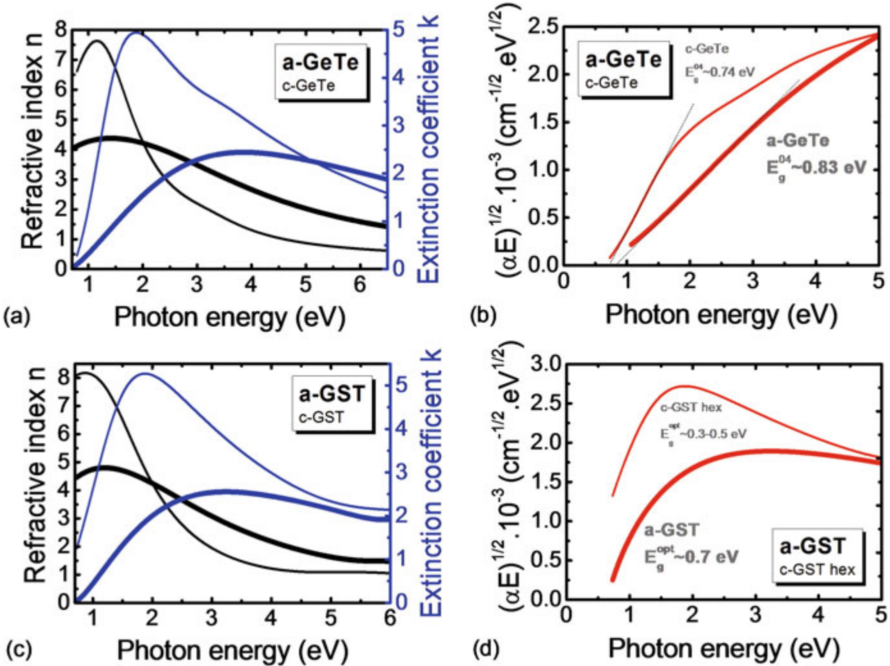


Fig. 6.25 (**a**) and (**c**): Optical constants (refractive index *n* and extinction coefficient k) extracted from spectroscopic ellipsometry measurements by using a Tauc-Lorentz law in the case of the amorphous state and with addition of two oscillators plus a Drude contribution (to account for the free carrier contribution) in the crystalline state. (**b**) and (**d**): Absorption $(\alpha.E)^{1/r}$ (with $r = 2$) versus photon energy E. The coefficient r is chosen equal to 2 since amorphous and crystalline GeTe and $Ge_2Sb_2Te_5$ are reported as indirect bandgap semiconductors. The samples are uncapped 100 nm thick films obtained by magnetron sputtering measured in the as-deposited (amorphous) state (thick lines) and in the crystalline state (thin lines) after heating at 400 °C. The crystallized GeTe film is rhombohedral, and the crystallized $Ge_2Sb_2Te_5$ film is hexagonal

6.2.2.2 and 6.2.2.3), two main explanations, involving either Peierls distortion or resonant bonding, can be sketched as follows (see also [92, 143]).

In the amorphous phase, Ge(Sb) atoms are found to be either in a tetrahedral environment (corresponding to sp^3 bonding) or a (3 + 3) distorted octahedral environment (corresponding to p bonding). Only the second environment is present in rhombohedral GeTe and cubic GST phases. In crystals, the presence of alternate short and long Ge(Sb)-Te distances in octahedra (see Fig. 6.15) can be understood as the signature of a Peierls (symmetry breaking) distortion. By comparison with a hypothetical undistorted and perfectly cubic structure (corresponding to p bonding), in which Ge(Sb) occupy octahedron centers, the Peierls distortion reduces the electronic energy by opening a gap at the Fermi level. The larger the distortion, the larger the gap. Peierls distortion can also stabilize amorphous phases and even liquids by locally alternating shorter and longer bonds [41]. In the case of GeTe, $Ge_2Sb_2Te_5$, and other Ge-Sb-Te phases on the pseudo-binary line such as

$Ge_1Sb_2Te_4$, the difference between short and long Ge(Sb)-Te bonds in octahedra is larger in the amorphous phases than in the crystalline phases. As a result, the Peierls distortion is significantly larger in the amorphous phase compared to its crystalline counterpart [42, 92, 139]. A larger Peierls distortion would explain the larger bandgap in the amorphous phase compared to the crystalline phase and the large change in electronic properties upon crystallization.

Another explanation of the property contrast between amorphous and crystalline phase change materials is the existence of resonant bonding in the crystalline state of phase change materials [91, 92, 150] where each atom possesses three valence p electrons in average and has to be covalently bonded with six nearest neighbor atoms (considering that the s valence electrons do not participate in chemical bonding). As a result, resonance between a bonding state and a nonbonding state can occur, giving rise to so-called “resonance bonding,” as proposed by Pauling [127]. The resulting resonantly bonded crystal is characterized by a strong electron delocalization and a high electronic polarizability, leading to high optical dielectric constant and refractive index [150], in agreement with experiments. In the amorphous phase, the large structural distortions are expected to suppress resonant bonding. This idea is supported by the smaller optical dielectric constant measured in the amorphous state [150]. Note, however, that the existence of long and short bonds in the crystalline state should have a detrimental effect on resonant bonding. Further studies are required to clarify the bonding character of phase change materials.

6.2.4 Crystallization of Phase Change Materials

In this section, we first discuss briefly the current understanding of atomic rearrangements occurring during crystallization. Then, the classification of phase change materials in either “nucleation-dominated” or “growth-dominated” materials is sketched. Information on crystallization of phase change materials can also be found in Chaps. 3 and 4.

In a simplified approach, a fast crystallization suggests limited structural changes between the amorphous and crystalline state. A first model, referred to in the literature as the “umbrella flipping” model, was proposed in [74]. The authors focused on the local arrangement of Ge atoms in the amorphous and crystalline state, as deduced from EXAFS analysis. They concluded that the local arrangement of germanium atoms is tetrahedral in the amorphous state and that, at crystallization, Ge atoms flip from a tetrahedral position to an octahedral position. The reality turned out to be more complex. First, from our present understanding of the structure of amorphous and crystalline phases (Sects. 6.2.2.2 and 6.2.2.3), it is clear that tetrahedral and defective octahedral environments exist in the amorphous state, while only defective octahedral environments are found in the crystalline phases, and that octahedral environments are more distorted in the amorphous phase than in the crystalline phase. In addition Ge-Ge homopolar bonds are only

found in the amorphous phase. Recently, direct simulations of crystallization became possible, thanks to improvements in codes and in computing power [48, 63, 144, 154], and brought light on the structural changes occurring during phase transition, as shown, for instance, in Fig. 6.26.

In bulk amorphous materials, crystallization starts with the formation of crystalline nuclei inside the amorphous phase, followed by crystallite growth process. In mushroom-type PCM devices, the amorphous phase forms a small dome

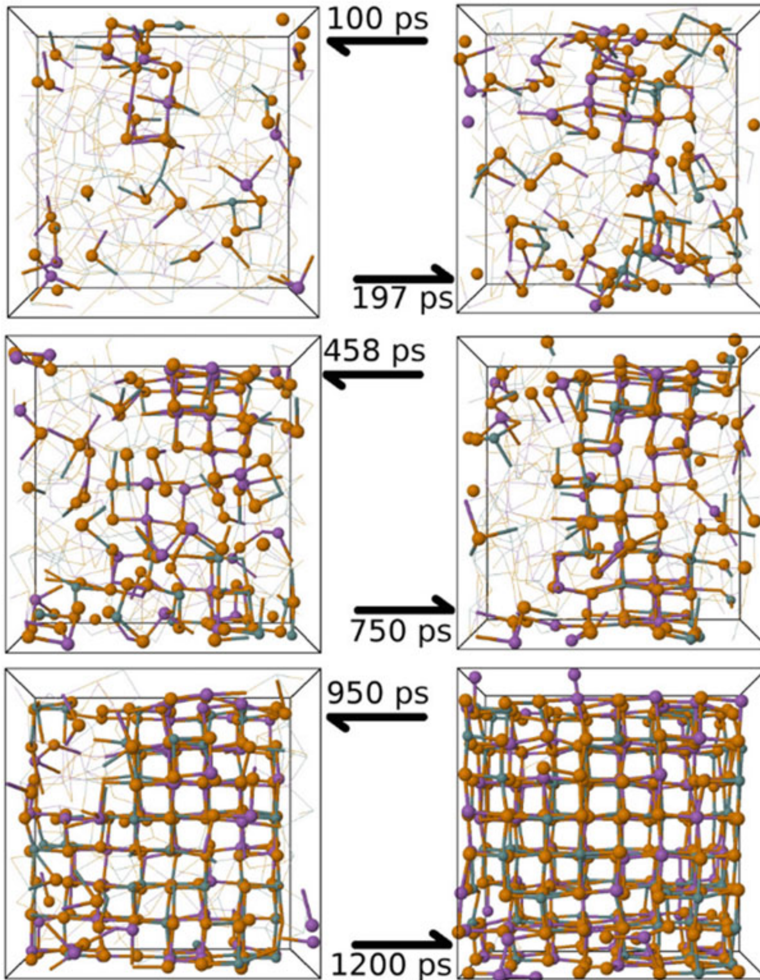


Fig. 6.26 Stages of crystallization of $Ge_2Sb_2Te_5$ simulated as a function of time (From [63]). Atoms in a crystalline environment are represented by spheres (green, Ge; purple, Sb; orange, Te). Bonds between atoms in the amorphous phase are represented by thin lines. *ABAB* square rings (*A*, Ge; Sb; *B*, Te) are present at the first stage (100 ps). Their number increases and they get more and more connected with increasing time

surrounded by the crystalline phase (see Fig. 6.6). In that case, crystallization can start either within the amorphous dome or at the amorphous/crystalline interface (another possibility being that crystallization starts at the interface with the bottom metal heater). Phase change materials have been divided into two classes [61, 105, 135]. In so-called "growth-dominated" materials, crystallization occurs preferentially by growth from the amorphous/crystalline interface. In "nucleation-dominated" materials, nuclei are formed inside the amorphous phase, this step being followed by a growth of the crystallites. $Ag_{5.5}In_{6.5}Sb_{59}Te_{29}$ (AIST), widely used in optical memories, is a prototypical example of a "growth-dominated" material, while $Ge_2Sb_2Te_5$ is reported as a "nucleation-dominated" material [61]. The difference between the two materials is illustrated in Fig. 6.27. In "nucleation-dominated" materials, the crystallization rate does not depend on the volume of the amorphous region. This is not the case for "growth-dominated" materials for which the crystallization rate, and thus the writing speed in devices, depends on size, shape, and nature of interfaces of the amorphous region.

The distinction between "growth"- and "nucleation-dominated" materials has been successfully used to optimize optical storage media (achieving short crystallization time and high data density in DVDs and Blu-ray discs). Crystallization kinetics in the temperature range reached in PCMs is currently under the scope of intense research efforts [55, 122, 123, 146, 148, 173], as illustrated in Fig. 6.27e. The current conclusion is that the nucleation rate plays the major role for controlling crystallization process at nanosecond timescale and nanometer length scale. Nevertheless, the reason why several phase change materials, such as GST alloys, exhibit a high nucleation rate whereas others, like AIST, have a low nucleation rate must be addressed. One must also explain the coexistence of a high stability of the amorphous state at low temperature and a fast crystallization. Bringing light on this issue requires the determination of the crystal growth velocity up to elevated temperature in the supercooled liquid. In the case of $Ge_2Sb_2Te_5$, the activation energy for crystal growth determined from ultrafast heating calorimetry was found to depend on temperature [122]. By opposite, many glasses, such as SiO_2, exhibit a temperature-independent activation energy for crystal growth with an Arrhenius behavior. Based on the concept of fragility, introduced by Angell [5, 6], the conclusion is that supercooled $Ge_2Sb_2Te_5$ has a fragile behavior [122] (see more information on this issue in Chap. 3).

Besides, from the study of the temperature dependence of viscosity, the differences between a "nucleation-dominated" material (GST alloys) and a "growth-dominated" one (AIST alloys) can be elucidated [123]. While GST exhibits a fragile behavior, AIST exhibits a fragile-to-strong crossover on cooling. This difference leads to different temperature dependence of the growth rate in the two materials and has direct consequences on PCM technology. In a PCM cell, the optimum switching operation would be obtained close to the maximum growth rate U. Whereas both phase change materials have similar maximum U values, these maxima occur at different temperatures: for AIST U is maximum at $T/T_m = 0.89$ (or $T_g/T = 0.52$), while for GST U is maximum at a significantly lower T/T_m value of about 0.76 [123]. Besides, the maximum rate of homogeneous nucleation occurs

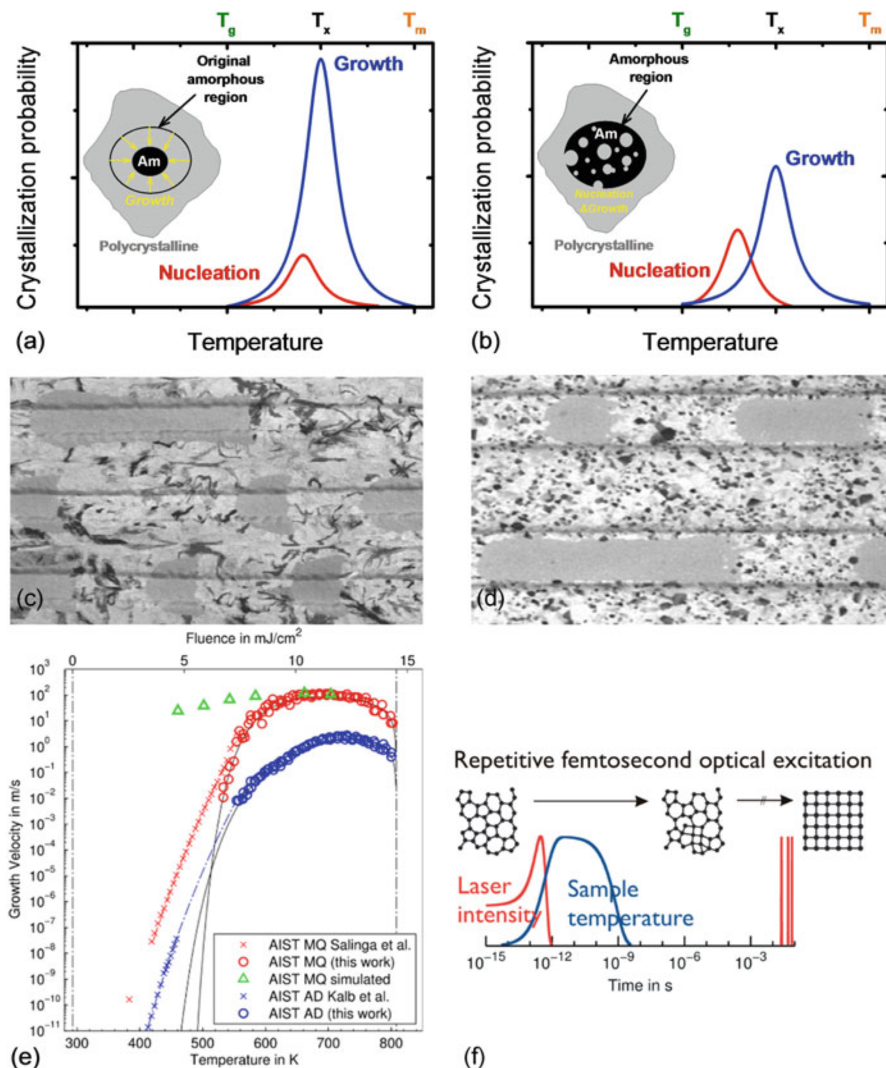


Fig. 6.27 (**a**) and (**b**) Schematic illustration of the probability of nucleation and growth as a function of temperature in the range between the glass transition temperature T_g and the melting temperature T_m for (**a**) a "growth-dominated" and (**b**) a "nucleation-dominated" material. The inserts in figures (**a**) and (**b**) illustrate the crystallization of an amorphous mark embedded in a crystalline layer. (**c**) and (**d**): Transmission electron microscopy (TEM) images of phase change layers for which the crystallization process is dominated by (**c**) growth and (**d**) nucleation, as evidenced by the resulting crystalline morphology (From [105]). (**e**) Temperature dependence of the crystal growth velocities of as-deposited (AD) and melt-quenched (MQ) AIST measured by monitoring the change in reflectance during femtosecond laser pump-probe experiments, as illustrated in (**f**) (From [173]. Copyright © 2015 American Chemical Society)

typically around $T/T_m \approx 0.6$ due to nucleation barrier arising from interfacial energy between crystals and liquid [67]. Based on the huge difference of U values at $T/T_m \approx 0.6$, one expects that the maximum nucleation rate is lower for AIST than for GST by about six orders of magnitude.

6.3 Stability and Aging of Phase Change Materials

6.3.1 Effect of Doping on the Structure of Phase Change Materials

Data retention is determined by the ability of a PCM cell to avoid unintended recrystallization (see Chap. 5). At a given temperature, the retention time is defined as the time elapsed before crystallization of the amorphous phase. Hence, increasing data retention requires to improve the stability of the amorphous phase against crystallization. For PCMs, the required data retention is 10 years at the operating temperature (≈ 150 °C in the case of embedded applications). The retention time as a function of the inverse of temperature obeys an Arrhenius law (see Fig. 7.7 in Chap. 7), which allows by extrapolation to determine the temperature at which it is equal to 10 years. The retention of $Ge_2Sb_2Te_5$ and GeTe (10 years at 75 °C and 110 °C for $Ge_2Sb_2Te_5$ and GeTe films, 87 °C and 101 °C in $Ge_2Sb_2Te_5$ and GeTe test devices) is not sufficient for embedded applications.

Thus, strategies have been developed to increase the crystallization temperature of phase change materials, among them inclusion of suitable dopants (such as B, C, N, O, SiC, SiN, or SiO_2 [10, 11, 35, 149, 176]) at relatively high concentration (a few percent) in Ge-Sb-Te and GeTe alloys. The most studied light dopant elements are C and N. Their effects on crystallization temperature T_x, activation energy for crystallization E_a, crystallization time, and resistivity contrast are extensively reviewed in Chap. 7. Hereafter, we investigate how dopants affect the structure of the amorphous state of phase change materials and can enhance its stability. Thanks to combination of ab initio molecular dynamics simulations and experimental studies of the local structure and dynamics of the amorphous phase [19, 25, 43, 60, 69, 79, 89, 140], these issues are fairly well understood. We will also discuss briefly the location of dopants in the crystallized state. In the following discussion, we focus on C-doped and N-doped GeTe. Undoped GeTe exhibits better retention properties than undoped $Ge_2Sb_2Te_5$. The same conclusion remains true in C-doped and N-doped alloys. Incorporation of 10 at % of C (resp. N) atoms in GeTe increases T_x by 135 °C (resp. 85 °C), and these significant increases of T_x are accompanied by a reduction of programming currents [10, 11, 128]. The activation energy for crystallization E_a also increases with doping (see Chap. 7). Doping GeTe with C or N is thus especially promising for applications. Besides, the fact that GeTe is a binary alloy facilitates ab initio molecular dynamics simulations and interpretation of experimental studies of the structure and dynamics. Hence, doped

GeTe is a case study to investigate the origin of the stabilization of the amorphous phase by dopants in phase change materials.

From ab initio simulations coupled with pair distribution function measurements on doped amorphous samples made by sputtering, it is found that C deeply modifies the structure of the amorphous GeTe phase through long carbon chains and tetrahedral or triangular units centered on carbon [43]. These units generate a new Ge-Ge distance ≈ 0.33 nm (absent in undoped GeTe). Besides, the characteristic distances in amorphous GeTe, namely, Ge-Te distances and short Ge-Ge distance at 0.247 nm (see Sect. 6.2.2.3), are almost unaffected by doping. The different types of carbon environment in amorphous C-doped GeTe are schematically represented in Fig. 6.28. In the case of amorphous N-doped GeTe, simulations [140] show that N is either bonded to Ge atoms, forming pyramidal NGe_3 and tetrahedral NGe_4 units, or found in molecular N_2, as illustrated in Fig. 6.28. Remarkably, in these simulations, no N-Te bonds are present, whereas C atoms are bound to Ge and Te. EXAFS experiments at the N K-edge in amorphous N-doped GeTe [79, 89] show that N is predominantly bonded to Ge atoms. It is unclear whether or not N_2 molecules are present in as-deposited N-doped GeTe films. N_2 molecules have been detected in [89] but not in [79]. Their concentration could depend on the details of the film deposition method. In the case of N-doped $Ge_2Sb_2Te_5$ amorphous phase, N is also

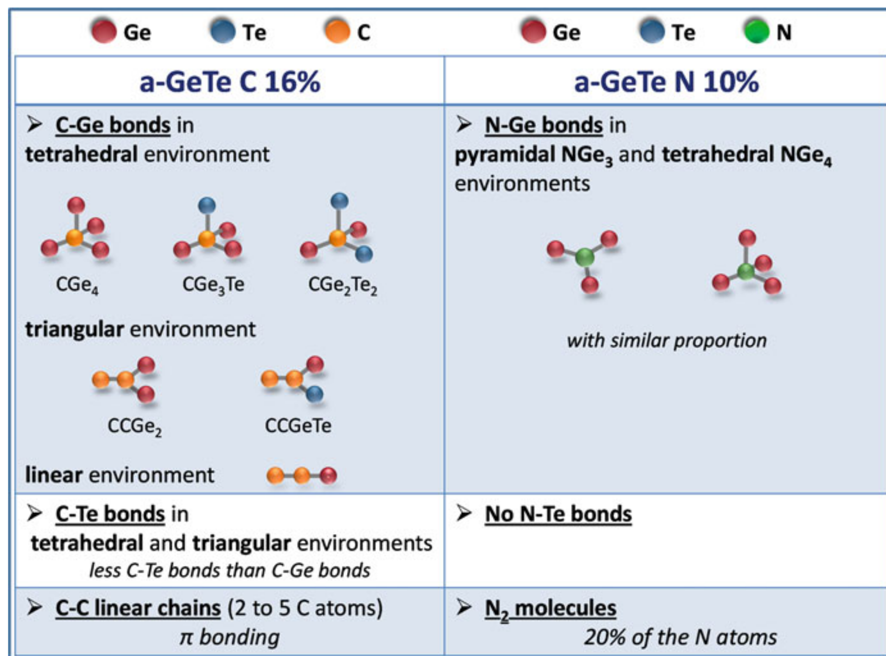


Fig. 6.28 Schematic drawing of the different structural units introduced by C and N atoms in the amorphous phase of C-doped and N-doped GeTe, as deduced from ab initio simulations [43, 140]

found to be bound to Ge and not to Te [69, 19, and references therein], and N_2 molecules have been detected experimentally [60].

Insight about amorphous phase stability can be gained by studying its dynamics, as previously shown in the general case of glasses [9, 24, 37, 113]. The FTIR absorbance spectra of undoped, C-doped, and N-doped GeTe samples are plotted in Fig. 6.29, as well as the calculated vibrational density of states (VDoS) [140]. Clear changes are observed upon doping. In the frequency range above 8 Thz, where undoped GeTe has no absorption, vibration modes involving dopants are present. Their intensity grows with dopant concentration. Thanks to calculations of the dopant site-projected VDoS, these modes can be related to the various environments presented in Fig. 6.28 [140]. In the range 3–7 THz, the density of vibration modes decreases upon doping, as observed in Fig. 6.29. Note that the same conclusion holds below 3 THz where only simulations are available. Thus, the weight of the vibration modes of undoped GeTe is attenuated by doping. In summary, high-frequency localized modes involving dopants are created, and the density of the low-frequency modes is depleted upon doping.

During the crystallization process, bonds have to be broken to create the new local environments of the crystalline structure. The level at which local atomic environments are mechanically constrained in the amorphous state thus influences its stability. In simulations, the degree of stress can be evaluated by using the Maxwell constraint counting method, as described in [107]. In the case of amorphous GeTe, the stress rigid character increases upon doping with C and N, essentially because of a slight increase of the tetrahedral character of the environment of Ge atoms [140]. This effect correlates with the increase of the crystallization temperature and activation energy for crystallization upon doping. Nevertheless, increase of mechanical rigidity of the amorphous phase upon doping is not the only relevant parameter to explain the increased stability. Changes in the dynamics in the low-frequency range must also be examined. Simulations show a reduction of the VDoS peak at 1.55 THz upon doping. This peak corresponds to the so-called boson peak that is characteristic of collective excitations in glasses [113]. In glasses, a decrease of the boson peak height has been correlated to an increased glass stability [37], and both theory and experiment have shown that the boson peak height decreases when the number of mechanical constraints increases [9, 37, 113]. Therefore, all changes of the dynamics induced by doping GeTe are related. In summary, studying the dynamics of doped amorphous GeTe brings light on the stabilizing effect of light dopants, such as C and N. The dopants create high-frequency localized modes. At the same time, they reduce the amplitude of the boson peak at low energy. They also increase the mechanical rigidity of the amorphous state. All these effects contribute to increase the stability of the amorphous state upon doping and lead to improved data retention in devices.

Having clarified the impact of dopants on the stability of the amorphous phase of GeTe, a relevant issue for applications can now be examined. What is the location of C and N doping atoms after crystallization of GeTe? For that purpose, one can study how the vibrational modes involving dopants, observed in the amorphous state in the frequency range above 8 Thz (Fig. 6.29), evolve during the

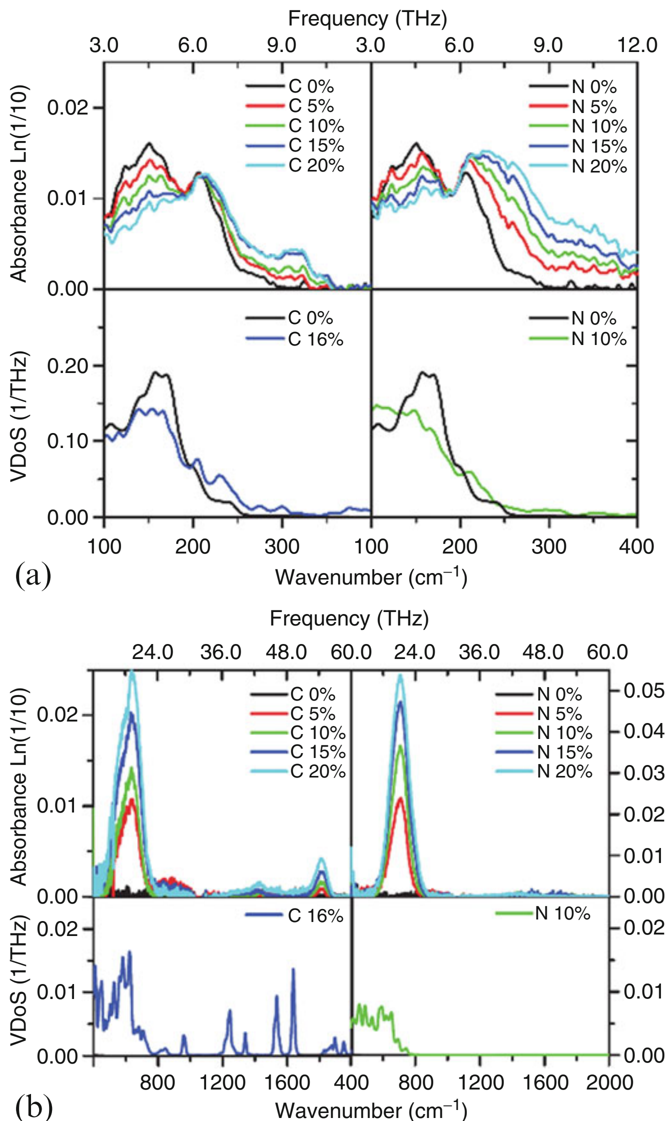


Fig. 6.29 (**a**) Experimental low-frequency FTIR absorbance spectra of as-deposited (amorphous) undoped, C-doped, and N-doped GeTe thin films (upper panels). Calculated VDoS are plotted in the lower panels. The FTIR spectrum and VDoS of undoped GeTe are null above 8 THz. (**b**) Experimental high-frequency FTIR absorbance spectra of as-deposited undoped, C-doped, and N-doped GeTe thin films (upper panels where different scales are used for C and N doping). The FTIR spectrum of undoped GeTe is null in the high-frequency range. Calculated VDoS are plotted in the lower panels. The simulations well reproduce the experiment trends, given that a direct comparison of the amplitudes is meaningless, the FTIR absorbance depending on the nature of vibration modes (From [140])

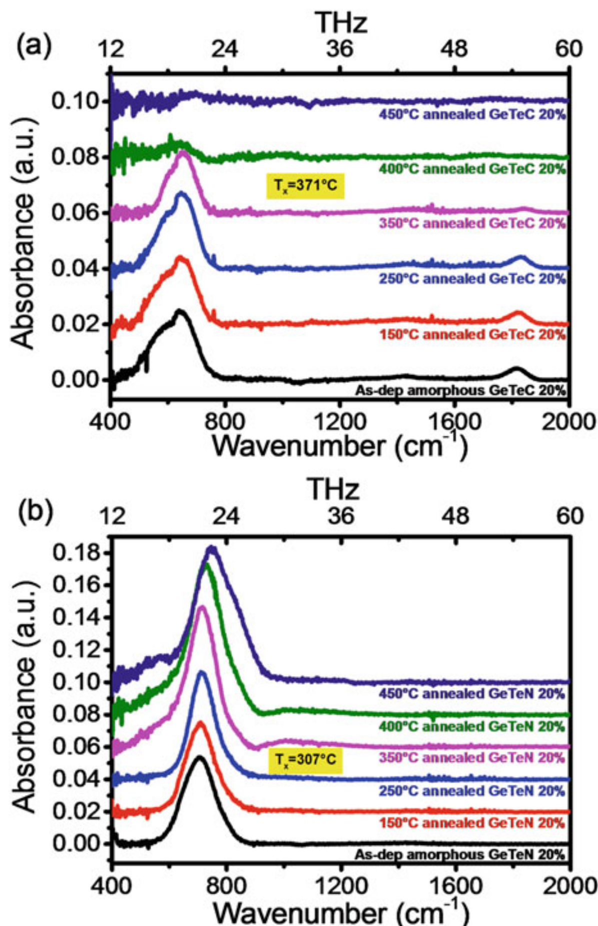


Fig. 6.30 Room temperature experimental high-frequency FTIR absorbance spectra of (**a**) a C-doped GeTe film (20 at % of C) and (**b**) a N-doped GeTe film (20 at % of N) as a function of annealing temperature. Crystallization occurs at T_x = 371 °C (**a**) and 307 °C (**b**)

crystallization of C- and N-doped GeTe films. Results are shown in Fig. 6.30. The high-frequency FTIR absorbance spectra change upon crystallization, but C-doped and N-doped GeTe films exhibit different behaviors. In C-doped GeTe, the high-frequency vibrational modes associated to -C-Ge or -C-Te bonds in the amorphous state disappear after crystallization. This result means that C atoms are not incorporated into the crystalline GeTe phase within the experimental limit of detection. In N-doped GeTe, the broad absorption peak around 700 cm^{-1} is attributed to NGe_3 and NGe_4 units [140]. As the annealing temperature increases above T_x, contrary to the case of C-doped GeTe, the absorbance does not vanish but progressively changes in the range [400–900 cm^{-1}], with, in particular, the appearance of two shoulders around 550 and 800 cm^{-1}. These spectra bear a close resemblance to those measured in amorphous germanium nitride (GeN_x) [174]. Changes in N-doped GeTe films upon heating above T_x can thus be interpreted by the progressive formation of a more and more N-rich amorphous GeN_x phase.

In crystallized C-doped GeTe films, the C atoms rejected from the crystalline GeTe phase must form a carbon phase, which is expected to be located at the grain boundaries of polycrystalline GeTe. This carbon segregation should limit the grain growth after crystallization. This is what is observed by X-ray diffraction studies, as shown in Fig. 6.31. The crystallite size measured at room temperature after heating to 450 °C is found to decrease as the C concentration increases (grain refinement). This result is consistent with that observed in C-doped $Ge_2Sb_2Te_5$ films (see, for instance, [177] where, in addition, amorphous carbon clusters could be observed in high-resolution transmission electron microscopy images of crystallized films). The localization of the GeN_x phase in crystallized N-doped GeTe films is not established. However, a reduction of the crystallite size is also observed as the N concentration increases (Fig. 6.31, see also [71, 128]). Thus, one can expect that the GeN_x phase is segregated at the GeTe grain boundaries. Similar conclusions are

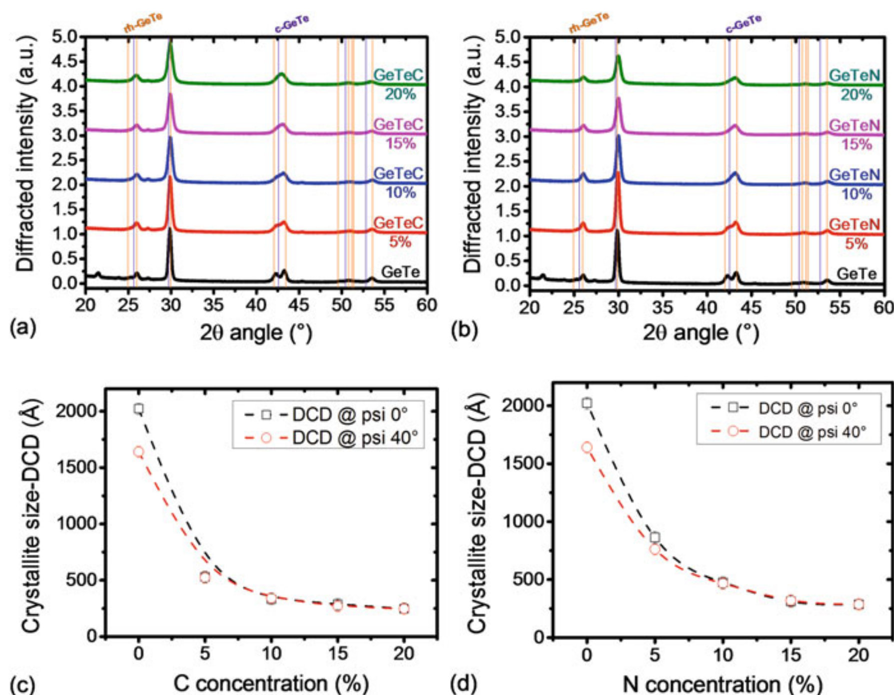


Fig. 6.31 Room temperature X-ray diffraction ($\lambda = 0.1542$ nm) pattern of (**a**) C- and (**b**) N-doped crystallized GeTe films with different C and N contents. All films have been previously heated to 450 °C. The theoretical positions of the diffraction peaks corresponding to rhombohedral (rh-GeTe) and cubic (c-GeTe) phases are indicated by vertical lines. The GeTe crystalline phase evolves from rhombohedral to cubic as the C or N doping level increases. (**c**) and (**d**): Change in crystallite size as a function of C or N concentration. The crystallite size was deduced by measuring the full width at half maximum (FWHM) of the (012) X-ray diffraction Bragg peak ($2\theta = 30°$) and by calculating the diffractive coherent domain (DCD) using Scherrer's law (Courtesy of F. Fillot)

reached in the case of N-doped $Ge_2Sb_2Te_5$ [58, 19, 160, and references therein]. In particular, the simulations in [19] show that insertion of N in the atomic form is energetically costly in the crystalline $Ge_2Sb_2Te_5$ phase, which explains why N atoms are expelled at crystallization. Note that a few N_2 molecules could remain inserted in crystalline $Ge_2Sb_2Te_5$ [19]. Finally, the effect of N on the crystallization kinetics of $Ge_2Sb_2Te_5$ has been recently studied by simulation [90].

In summary, C and N can be incorporated in the amorphous phase of GeTe and $Ge_2Sb_2Te_5$ but not in the crystalline phases (one exception could be the presence of a small number of N_2 molecules in the crystalline state of N-doped alloys). In the crystallized state, dopants are rejected in an amorphous phase located at the grain boundaries of the GeTe or $Ge_2Sb_2Te_5$ crystalline phase, which leads to grain refinement. This segregation could explain the observed decrease of thermal conductivity upon doping in C- and N-doped GeTe [33, 85]. A difference between C and N doping in GeTe (and probably in $Ge_2Sb_2Te_5$ as well) must be emphasized. Rejection of C does not change the Ge/Te ratio, while formation of a GeN_x amorphous phase in crystallized N-doped GeTe would lead to a Ge depletion of the crystalline GeTe phase. This observation brings light on an experimental observation in PCM devices. The beneficial effect of N doping on retention in GeTe-based PCMs is optimal around 2 at % of N [35]. This optimal N concentration is comparable to the slight Ge excess generally obtained in sputtered amorphous GeTe, which differs from the composition of the crystalline GeTe phase (see Sect. 6.2.2.2).

6.3.2 Structural Relaxation and Resistance Drift

One of the most promising properties of PCM technology is the multilevel cell (MLC) operation (see Chap. 11), which enables cost reduction and storage density increase. MLC takes the opportunity to store several bits per cell, by encoding them in intermediate resistance levels obtained from partial crystallization of the phase change material [94]. However, the resistivity of the amorphous phase increases with time (so-called “resistance drift phenomenon”; see Chap. 2) due to material aging. Drift can lead to memory failure and hindered up to now the development of multilevel storage devices. Thus, reducing the drift has become one of the major challenges for the development of PCMs. Note that drift reduction can be conflicting with other requirements. For instance, increasing Ge content in Ge-Sb-Te alloys increases the crystallization temperature T_x, but this change is accompanied by an increase of the resistance drift [180]. Understanding the origin of drift is thus required for material optimization. In literature, resistance drift is usually ascribed to a “structural relaxation,” but very different physical mechanisms have been proposed. In this section, after a description of the resistance drift phenomenon, we review the current understanding of structural relaxation of the amorphous phase during aging (for a review focused on electrical models of drift, see Chap. 2).

Then, we add a contribution to the current debate by discussing an EXAFS study of drifted amorphous GeTe films [115].

At a fixed temperature, the resistance drift of the amorphous phase leads to a progressive increase of the resistance of the RESET state in a PCM cell. It must be noted that resistance drift with aging is negligible in the SET state when the phase change material is in the crystalline state. Resistance drift in the amorphous state can also be measured on amorphous films, as illustrated in Fig. 6.32. The GeTe film was first heated from room temperature to the targeted annealing temperature at a heating rate of 10 °C/min, and the temperature was then maintained constant during several hours. During heating, the resistivity decreases due to the semiconducting behavior of amorphous GeTe. Then, at each constant temperature, a clear increase of the resistance (drift) is observed as a function of the elapsed time. Note that, at 80 °C, the drift starts only after an incubation time. The existence of an incubation time that depends on thermal history and also on aging after fabrication has been previously reported in [163]. In PCM devices, drift measurements are generally made soon after the RESET pulse, and thus no incubation occurs.

In literature, it has been proposed that, at a given temperature, the resistance change as a function of time obeys a power law (of the type $R(t) = R(t_0).(t/t_0)^{\nu}$) and that the exponent ν can be used to quantify the drift phenomenon [94]. However, it

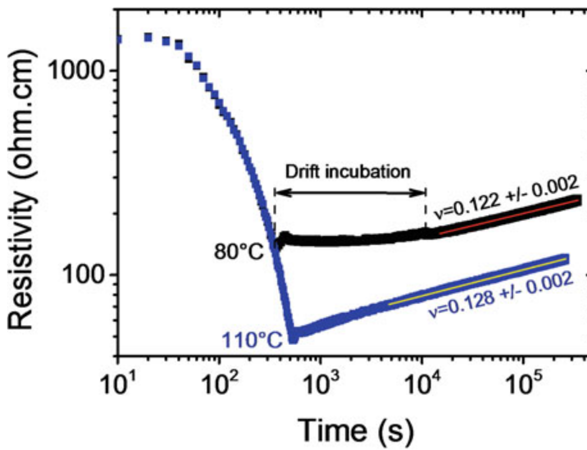


Fig. 6.32 Resistivity (logarithmic scale) measured by a 4-point probe as a function of time on a 100 nm thick amorphous GeTe film. The GeTe film has been deposited on a 500 nm thermal SiO_2 layer on top of a Si substrate. It has been protected from surface oxidation by in situ deposition of a 10 nm SiN layer and thus crystallizes near 230 °C (see Sect. 6.4.2). Pieces of this film were first heated at 10 °C/min from room temperature to either 80 °C or 110 °C. When the temperature increases, the resistivity decreases since amorphous GeTe has a semiconducting behavior. Then, the temperature was maintained constant (annealing process) for 48 h at 80 °C or 72 h at 110 °C. Drift exponents ν of 0.122 at 80 °C and 0.128 at 110 °C were obtained from fits in restricted time range (see text). At 80 °C, drift arises only after an "incubation" time. In the present case, the GeTe film was kept 8 months at room temperature after deposition before measurement, so that a significant waiting time is required to observe the resistivity drift at 80 °C

	$Ge_2Sb_2Te_5$	GeTe	Doped Sb_2Te
This work	0.099 ± 0.003 (ad)	0.123 ± 0.004 (ad)	0.055 ± 0.002 (ad)
Boniardi [16, 31, 41, 42]	0.09–0.11 (mq)		
Cho [45]	0.076 (ad)		
Ciocchini [46]	0.115–0.119 (mq)		
Fantini [34]	0.076–0.118 (mq)[1]		
Ielmini [47]	0.077 (mq)		
Ielmini [26, 27]	0.04–0.11 (mq)[2]		
Karpov [29, 33]	0.03–0.10 (mq)[3]		
Krebs [17]	0.138 ± 0.002 (ad)	0.122 ± 0.002 (ad)	
Luckas [44]	0.138 ± 0.005 (ad)	0.129 ± 0.005 (ad)	0.059 ± 0.005 (ad)
Mitra [35]	0.002–0.086 (mq)[3,4]		
Oosthoek [32]			0.081 (mq)
Papandreou [10]	0.018–0.06 (mq)[3]		
Pirovano [12]	0.06 (mq)		
Rizzi [43]	0.092–0.107 (mq)		

Fig. 6.33 Drift exponents for amorphous as-deposited (ad) and melt-quenched (mq) phase change materials from [163]. Note that the ν value for $Ge_2Sb_2Te_5$ ($\approx$0.1) is slightly smaller than for GeTe (0.123)

must be emphasized that the time dependence of the resistance is more complex and cannot be characterized by a single value of ν. Power law fits can only be performed in restricted time ranges. The measured ν values depend on the fitted time window, and besides they can be affected by the age or the thermal history of the thin film sample before resistance measurement. Thus, comparison of drift exponents between different phase change materials may be delicate. Nevertheless, it is clear that the drift varies with material composition, as illustrated in Fig. 6.33.

The underlying physical mechanisms at the origin of the drift phenomenon are still unclear and are currently highly debated. Different and sometimes opposite mechanisms have been proposed, such as:

- Increase [36] or decrease [53, 163] of disorder during drift
- Increase [130] or decrease [52] of the defect density of states during drift
- Drift impacted [108] or not impacted [145] by stress

The only consensus that has been reached in literature is that drift is due to a kind of relaxation of the amorphous structure with aging. This confused situation resulted mainly from the lack of structural studies and ab initio simulations of the amorphous phase as a function of aging. Some progresses have been made recently since several studies have addressed this issue [97, 109, 40, 141, 115, 179]. However, they lead to contradictory conclusions, and the debate is still going.

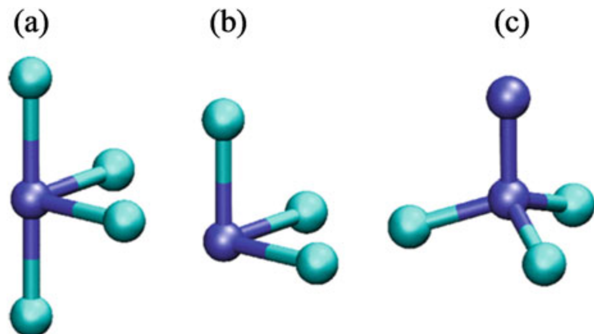


Fig. 6.34 Schematic drawing of the structural units present in amorphous GeTe. Ge atoms are found in (**a**) defective octahedra, (**b**) pyramids, and (**c**) tetrahedra with a Ge-Ge bond (From [40])

In the following, we focus on amorphous GeTe. This alloy exhibits a high drift coefficient $\nu \sim 0.123$ and can be considered as a prototypical case for structural studies on the origin of drift since it has already been the subject of numerous ab initio molecular dynamics simulations and advanced structural characterizations (see Sect. 6.2.2.3). Let us recall here that the local environment of Ge in amorphous GeTe is a mix of tetrahedrally bound Ge and Ge in defective octahedral environments with bond angles close to 90°, as illustrated in Fig. 6.34 from [40]. In distorted octahedral sites, Ge is mainly surrounded by Te, Ge-Ge bonds being mainly present in the tetrahedra. The number of Ge-Ge bonds has thus been directly related to the number of Ge in tetrahedral environments. Furthermore, from energy calculations, Ge-Ge bonds are proposed to be responsible for the stabilization of tetrahedrally coordinated Ge [27, 141].

The relaxation of the amorphous state and the resistance drift have been ascribed to the removal of tetrahedrally coordinated Ge sites, and thus of Ge-Ge bonds, leading to an increase in the bandgap and the removal of localized defects in the gap [40, 141]. A different mechanism is invoked in [80]. In this model, Ge-Ge bonds in amorphous GeTe generate free electrons that contribute to the material conductivity, and resistance drift is ascribed to the breaking of Ge-Ge bonds and to the resulting decrease of the number of free charge carriers. Other models have been proposed since many different types of structural defects [53] (e.g., dangling bonds, distorted bonds, vacancies, etc.) can also introduce localized states in the bandgap and evolve with aging. Besides, an alternative scenario was recently proposed from atomistic simulations [179]. Resistance drift would be caused by consumption of Ge atom clusters in which the coordination of at least one Ge atom differs from that of the crystalline phase and by removal of stretched bonds in the amorphous network, leading to a shift of the Fermi level toward the middle of the bandgap.

All these theoretical models involve changes of the local environment of Ge atoms and a reduction of the number of Ge-Ge bonds during drift. To get evidence of such structural changes, one can probe the local structure around Ge atoms by X-ray absorption spectroscopy at the Ge K-edge on amorphous GeTe thin film samples previously set at different resistance drift levels by thermal annealing. Results of such a study [115] are shown in Fig. 6.35. As the resistance drift increases (from sample D0 to sample D3), the intensity of the Fourier transforms

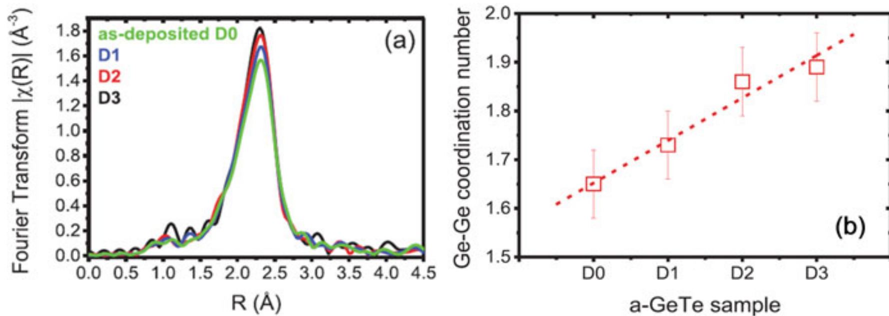


Fig. 6.35 (**a**) Fourier transforms of EXAFS (extended X-ray absorption fine structure) spectra measured at room temperature on four amorphous GeTe samples set at different aging and resistance drift states by different thermal annealing treatments. The resistance drift increases from sample D0 to sample D3. The sample denoted D0 corresponds to the as-deposited film and is shown here as a reference. The three other samples have been obtained after annealing for 1 (D1), 15 (D2), and 30 (D3) min at 180 °C, i.e., 50 °C below the crystallization temperature (230 °C) of the GeTe film, which has been capped in situ with SiN (see Sect. 6.4.2). (**b**) Change of the Ge-Ge coordination number deduced from fits of the EXAFS data with the resistance drift level. The dashed line is only a guide for the eye (Figures are reproduced from [115])

peak at 2.3 Å increases. This observation allows to conclude unambiguously that the aging of an amorphous GeTe film is accompanied by changes of the local structure around Ge atoms. The measured EXAFS data can be perfectly modeled with two coordination shells Ge-Ge and Ge-Te. The mean Ge-Ge distance ($R_{Ge\text{-}Ge} \approx 0.247$ nm) and the mean Ge-Te distance ($R_{Ge\text{-}Te} \approx 0.262$ nm) are identical, within experimental accuracy, in all samples and are thus insensitive to drift. These distance values are in good agreement with literature [98, 75, 32,]. Strikingly, $N_{Ge\text{-}Ge}$, the Ge-Ge coordination number, is found to increase with aging and drift of the amorphous GeTe film (Fig. 6.35b). No change of $N_{Ge\text{-}Te}$, the Ge-Te coordination number, is detected within the error bars. The width of the distance distributions is also found independent of drift. The increase of the Fourier transform with drift can thus be ascribed to an increase of the number of Ge-Ge bonds. According to literature where the number of Ge-Ge bonds has been related to the number of Ge in tetrahedral environments, this observation would mean that the number of tetrahedra in amorphous GeTe increases with aging. By contrast, in ab initio simulations [141], the removal of tetrahedrally coordinated Ge atoms was found to reduce the energy of the amorphous state. Hence, a decrease of Ge-Ge bonds would have been expected during structural relaxation. The same trend is predicted in [40].

One possible explanation of the EXAFS results would be a clustering of Ge atoms in the amorphous phase of GeTe during aging. The presence of Ge clusters can explain the detected increase of the number of Ge-Ge bonds and would introduce new states in the electron density of states (DoS) of amorphous GeTe, hence increasing its resistivity. Formation of Ge nanoclusters has not been detected in ab initio simulations. One reason could be that the timescale required for their formation (which involves atomic diffusion) is not accessible to ab initio

simulations. In experiments drift times are of the order of several hours. In conclusion, despite all these recent research efforts, no clear picture of the structural relaxation mechanism occurring during aging of the amorphous phase of phase change materials exists yet, and further investigations are needed.

6.4 Impact of Confinement on Phase Change Materials

6.4.1 Issues

Reduction of the programming currents in PCM cells is a major issue (see Chap. 3). This can be achieved by optimizing material composition, for instance by doping [50], as shown in Chap. 7. Another possibility consists in changing the memory cell architecture. Incorporation of the phase change material in a small confined structure allows to significantly improve the efficiency of Joule heating and reduce the RESET current with respect to the case of a usual planar mushroom cell [64, 70]. However, in confined structures, the phase change material is in contact with insulators (SiO_2, SiN, etc.) in addition to metallic electrodes (Ta, TiN, W, etc.), which raises new issues about impact of geometrical confinement on the material properties and phase transformations. In this context, it is mandatory to understand and to master interface effects due to encapsulating materials.

Numerous studies have been conducted to probe interface and size effects on phase change materials. Crystallization has been observed in very thin films (a few nm thick) [119, 133, 152, 161] as well as in small clusters [16, 44]. Until recently [45, 116], no influence of encapsulating materials on crystallization of phase change material films had been reported for thickness above 30 nm. In 30 nm thick $Ge_1Sb_2Te_4$ films, an influence of the cladding layer (SiO_2, Si_3N_4, Ta_2O_5, ZnS, or ZnS-SiO_2) on the crystallization process had been evidenced, but crystallization temperature (T_x) changes do not exceed ± 6 °C with respect to the SiO_2 reference case, a slightly larger effect (+13 °C) being detected only in the case of a Si_3N_4 cladding layer [119]. By contrast, as shown in Fig. 6.36, strong effects have been reported in $Ge_2Sb_2Te_5$ films with thickness less than 10 nm, with a spectacular increase of T_x (of 225 °C for a 2 nm thick film deposited on a Si substrate and capped with Al_2O_3 [133] and of 150 °C for a film sandwiched between TiN layers [152]). In the latter case, the large T_x increase has been ascribed to a compressive stress exerted by the TiN encapsulation layers [152]. In the case of ZnS-SiO_2 cladding, only small T_x changes are observed for film thickness under 10 nm, with conflicting results: decrease [152] and increase [161] of T_x when the thickness decreases.

From these studies, a consensus emerged that cladding materials significantly affect the crystallization of phase change material films only for thickness under 10 nm. It must be emphasized that these studies of scaling effects in thin films did not address the impact on crystallization of a possible interdiffusion of elements at interfaces between the phase change material and cladding layers. However, as

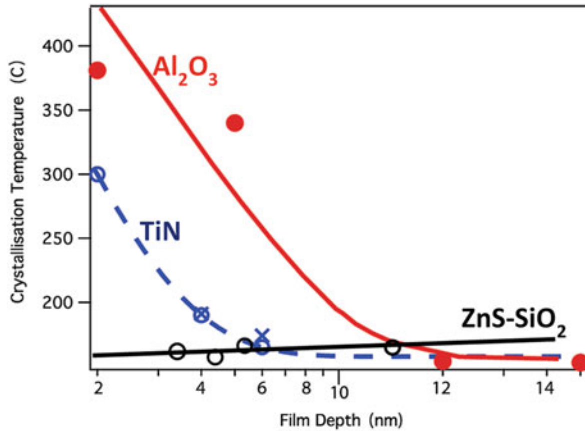


Fig. 6.36 Crystallization temperature of $Ge_2Sb_2Te_5$ films as a function of the film thickness and nature of cladding layers (From [152]. Copyright © 2009 American Chemical Society). Blue circles and crosses: films sandwiched with 20 nm thick TiN layers. Black circles: films sandwiched with 20 nm thick $(ZnS)_{0.85}(SiO)_{0.15}$ layers. Red dots: films deposited on a Si substrate and capped with Al_2O_3 from [133]

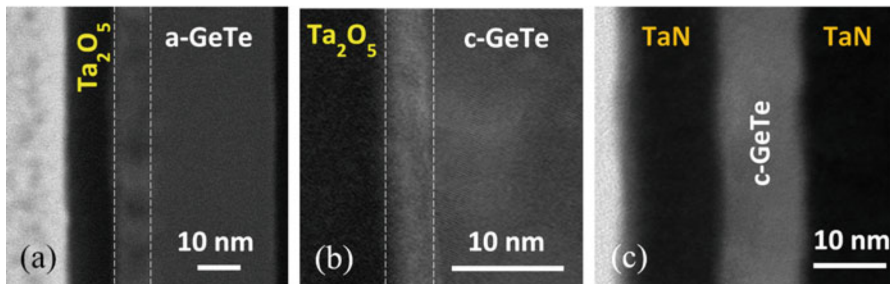


Fig. 6.37 Scanning transmission electron microscopy (STEM) images of GeTe films. (**a**) Amorphous as-deposited and (**b**) crystallized GeTe film sandwiched between Ta_2O_5 layers. Interdiffusion is detected at the top interface between GeTe and Ta_2O_5. (**c**) Crystallized GeTe film sandwiched between TaN layers. No interdiffusion is detected even after crystallization. These images evidence the need of a careful study of interfaces between encapsulating layers and the phase change material film prior to studying the scaling/interface effects on the phase change material crystallization

illustrated in Fig. 6.37, interdiffusion layers can exist at interfaces. Depending on the nature of the cladding material, their thickness varies from a few Å to a few nm. This observation raises issues about literature measurements on films with thickness in the range [2–5 nm].

Besides, we recently observed that the crystallization temperature of 100 nm thick $Ge_2Sb_2Te_5$ and GeTe films depends strongly on the cladding material [116] (Fig. 6.38). This effect that had previously gone unnoticed is due to a direct influence of the phase change material interface on the crystallization mechanism. Therefore, to bring light on size reduction effects on crystallization of phase change

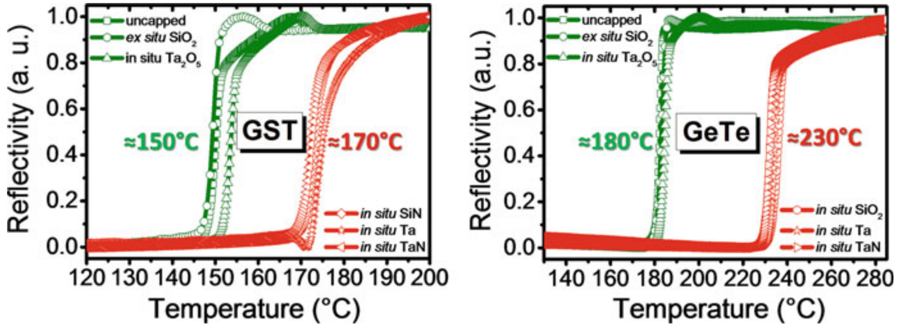


Fig. 6.38 Reflectivity at 670 nm as a function of increasing temperature (heating rate of 10 °C/min) for 100 nm thick $Ge_2Sb_2Te_5$ (labeled GST) and GeTe films either uncapped or capped with different top layers. "In situ" in the sample label means that capping was made by sputtering in the same high vacuum sputtering equipment as the phase change material film without vacuum break, i.e., without exposing the upper surface of the phase change material to air. To facilitate their comparison, the curves are normalized by setting the minimum value of the reflectivity to 0 and its maximum to 1. For each chalcogenide alloy ($Ge_2Sb_2Te_5$ or GeTe), the films fall into two groups: one in which the crystallization temperature T_x (~150 °C for $Ge_2Sb_2Te_5$ and ~180 °C for GeTe) is in agreement with usual literature values and another one with significantly higher T_x values (~170 °C for $Ge_2Sb_2Te_5$ and ~230 °C for GeTe)

materials, it is mandatory to first study interface effects in relatively thick films (≈ 100 nm thick) before studying either clusters or ultrathin films where size effects and interface effects are mixed. Interface effects will be discussed in Sect. 6.4.2. Scaling effects are discussed in Chap. 8. In the perspective of technological applications, an important issue is whether the crystallization mechanism of phase change materials can be controlled by interface effects. The answer is undoubtedly positive [116]. Through interface engineering, it is possible to select whether crystallization starts from the interface or inside the material.

6.4.2 Impact of Interfaces on Crystallization Temperature and Crystallization Mechanism of Phase Change Material Thin Films

Studying the influence of capping materials in 100 nm thick films allows to focus on interface effects [116]. For GeTe and $Ge_2Sb_2Te_5$ films deposited by sputtering and either left uncapped, and hence submitted to natural oxidation, or covered by a 10 nm thick capping layer (SiO_2, Ta_2O_5, Ta, TaN, or SiN), strong differences of crystallization temperature are detected from optical reflectivity measurements (Fig. 6.38). A remarkable trend, common to $Ge_2Sb_2Te_5$ and GeTe films, is the existence of two sample groups: one with crystallization temperatures corresponding to the values previously reported in the literature (~150 °C for $Ge_2Sb_2Te_5$ and ~180 °C for GeTe) and one with higher crystallization temperatures (~170 °C for $Ge_2Sb_2Te_5$ and ~230 °C for GeTe). Strikingly, the distribution of the

films in one group or the other is only decided by the capping and is the same for GeTe and $Ge_2Sb_2Te_5$.

The analysis of interfaces between $Ge_2Sb_2Te_5$ or GeTe and their capping layers as a function of their T_x group shows that all films belonging to the low T_x group are surface oxidized. It must be emphasized here that an "in situ" capping does not systematically ensure the absence of oxidation. Whether or not the chalcogenide phase change material surface has been exposed to O_2 during the deposition of the capping layer is the key parameter. The crystallization mechanisms involved are different in films with low or high T_x, as illustrated in Fig. 6.39 in the case of $Ge_2Sb_2Te_5$ where scanning transmission electron microscopy (STEM) images obtained at room temperature after heating to selected temperatures are compared [116]. In the case of an uncapped, and hence oxidized, 100 nm thick $Ge_2Sb_2Te_5$ film, the crystallization proceeds in two steps. The first one corresponds to the reflectivity increase at 151 °C and is associated to a first resistance drop. It involves surface nucleation at the upper oxidized surface of the $Ge_2Sb_2Te_5$ film. After this step, only the upper part of the film (on about 15–20 nm) is crystallized, the rest of the film being still amorphous. In films heated to a higher temperature, the only change is a slight growth of the crystallized part, and this until the heating temperature reaches 168 °C, which corresponds to the onset of a second resistance drop. Then, nuclei appear in the remaining amorphous part of the film. After heating to 175 °C, after the second resistance drop, the film is entirely crystallized. Hence, this second crystallization step involves nucleation inside the amorphous phase and growth. Partial crystallization of $Ge_2Sb_2Te_5$ films (corresponding to the first step reported here) had previously been observed [57, 62, 86, 159], but the full scenario of crystallization was not elucidated in these studies. Note that in reflectivity measurements, only the first crystallization step is detected because of the limited depth probed. By contrast, in a non-oxidized $Ge_2Sb_2Te_5$ film, crystallization occurs in one step and starts by nucleation inside the film volume. Only one resistance drop is observed (around 173 °C). It coincides with the reflectivity increase. It must be emphasized that, coherently, the one-step crystallization in non-oxidized film and the second crystallization step in oxidized films, which both involve volume nucleation, occur at the same temperature ($\approx$170 °C).

In summary, in $Ge_2Sb_2Te_5$ films, in the absence of oxidation, nucleation occurs inside the volume of the film, while in oxidized films, heterogeneous nucleation occurs at the oxidized upper surface. The same conclusions apply to GeTe. In addition, by controlling the degree of oxidation of the upper surface, one can control the onset of crystallization. This is illustrated in Fig. 6.40 in the case of $Ge_2Sb_2Te_5$. When exposing the film upper surface to air for increasing durations before capping, the temperature of the crystallization onset progressively decreases.

In most studies in the literature, information on the deposition method of the capping layer is insufficient to determine whether or not the phase change material surface has been oxidized during fabrication. Moreover, oxidation after deposition and during resistance or reflectivity measurements could occur in the case of too thin or porous capping layers. It must be noticed here that onset of crystallization at a temperature ~ 150 °C for $Ge_2Sb_2Te_5$ and ~180 °C for GeTe can be taken as the

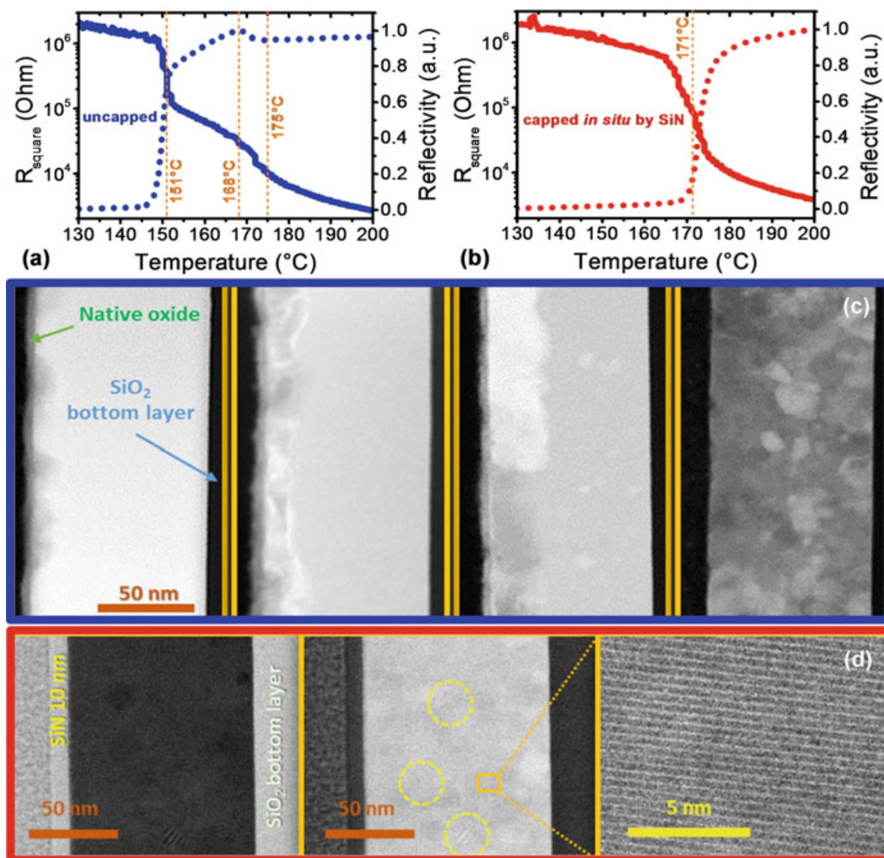


Fig. 6.39 (**a**) Uncapped, and hence oxidized, 100 nm thick $Ge_2Sb_2Te_5$ film. On left, resistance (continuous line) and reflectivity at 670 nm (dashed line) as a function of increasing temperature (heating rate of 10 °C/min). Two resistance drops are detected. The first one coincides with the reflectivity increase. (**b**) 100 nm thick $Ge_2Sb_2Te_5$ film protected from surface oxidation by in situ deposition of a SiN capping layer. On left, resistance (continuous line) and reflectivity at 670 nm (dashed line) as a function of increasing temperature (heating rate of 10 °C/min). The resistance decreases in one step, which coincides with the reflectivity increase. (**c**) Room temperature scanning transmission electron microscopy (STEM) images of the oxidized film in different states. From left to right, amorphous as-deposited state, after heating to 151 °C (end of the first resistance drop in (a)), after heating to 168 °C (at the beginning of the second resistance drop), and after heating to 175 °C. A two-step crystallization process is evidenced (see text). (**d**) Room temperature STEM images of the non-oxidized film. Left image: amorphous as-deposited state. Middle image: after heating to 171 °C (corresponding to resistance and reflectivity changes in (b)), small nuclei are detected in the volume of the film, highlighted by dashed circles on the figure. Right image: a zoom on one on these nuclei allows to evidence crystalline planes

signature of an oxidized phase change material surface. Finally, studies on crystal growth velocity in thin films have to be revisited since, in most cases, the measured velocity characterizes crystal growth in the plane of the film. In oxidized films, this

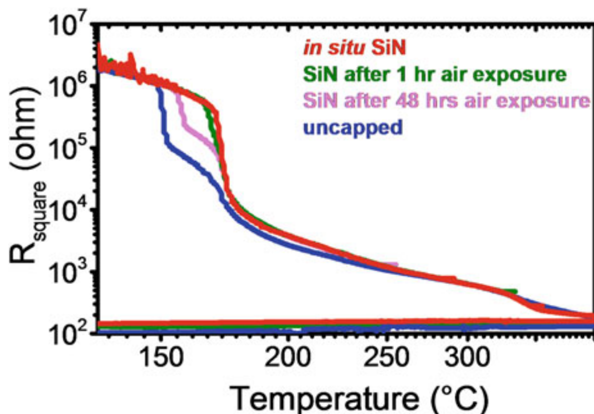


Fig. 6.40 Resistance as a function of temperature (heating and cooling rate of 10 °C/min) in 100 nm thick $Ge_2Sb_2Te_5$ films either left uncapped, capped in situ with SiN, or capped with SiN after exposure to air for either 1 h or 48 h. Upon heating, striking differences among the samples are observed at crystallization. The temperature of the onset of crystallization decreases from 173 °C in the non-oxidized sample to 151 °C in the uncapped sample that has been left 1 month in air after deposition. The resistance change around 330 °C, similar in all samples, is due to transformation of the crystalline fcc phase into the metallic hexagonal phase, which is then retained during cooling

velocity is much higher than the growth velocity in the direction perpendicular to the film surface and besides differs probably from the growth velocity after nucleation inside the amorphous phase.

The fact that heterogeneous nucleation at an oxidized surface occurs at a lower temperature than nucleation inside the volume is easily understandable if the latter nucleation is homogeneous. The nucleation barrier in a uniform substance is expected to be much higher than at a surface or on defects [96]. In view of the large difference of T_x (50 °C) between oxidized and non-oxidized GeTe films, compared to 20 °C in $Ge_2Sb_2Te_5$, homogeneous nucleation would be more difficult in GeTe than in $Ge_2Sb_2Te_5$. Why surface oxidation of chalcogenide phase change materials favors heterogeneous nucleation is still a pending issue. An oxygen doping effect can be excluded since oxygen, like other dopants, is known to increase the crystallization temperature [131, 178]. Interface energy effect due to the capping layer cannot be the leading cause since identical T_x are observed for all capping layers, provided that there is no oxidation of the phase change material surface. Heterogeneous nucleation at the material surface could result from the fact that oxidation of chalcogenide phase change materials is element selective [47, 171]. Ge oxidation occurs first. Thus, oxidation changes locally the phase change material composition, leading to the formation of Te-rich regions under the film surface [116]. These regions could crystallize at a lower temperature than the bulk material and act as preferential nucleation sites. Effects of surface oxidation on crystallization have previously been reported in other systems, such as ZBLAN glasses, with the selective formation of ZrO_2 [110] and metallic glasses in [82] and [124].

In conclusion, interface engineering allows one to select the crystallization mechanism of phase change materials such as GeTe and $Ge_2Sb_2Te_5$ alloys. In the absence of oxidation, heterogeneous nucleation at the material surface is eliminated, and nucleation occurs inside the volume of the phase change material. As a result, the crystallization temperature is higher in non-oxidized film, and the stability of the amorphous phase of $Ge_2Sb_2Te_5$ and GeTe can hence be increased by avoiding surface oxidation. This opens the route toward new PCM cell architecture with improved data retention, provided that no oxidation of the phase change material occurs during the fabrication process.

Starting from our present understanding of interface effects in relatively thick films, one can sketch several material issues about size reduction (PCM scaling is discussed in Chap. 8). Size reduction effects have been mainly explored on thin films of decreasing thickness (down to a few nm) [119, 133, 152, 161]. The observed changes of T_x when the thickness decreases below 10 nm have been ascribed to interface energy and/or stress effects, but the role of surface oxidation has been overlooked. Conclusions of these studies should thus be revisited. In particular, the increase of T_x with decreasing thickness has often been interpreted using a model [119] that assumes bulk nucleation and is not relevant in oxidized samples. In the literature on scaling effects, information given on the film deposition process is generally not precise enough to determine the oxidation level of the film. However, as shown above, a T_x value close to 150 °C in thick (above 30 nm) $Ge_2Sb_2Te_5$ films and close to 180 °C in thick GeTe films is the signature of oxidized films. This is the case for the results presented in Fig. 6.36. In an oxidized film, changes due to oxidation extend on a few nm under the upper surface (see Figs. 6.37 and 6.39a), so that results obtained on oxidized films with thickness below 5 nm are not relevant. Thus, the impact of scaling in thin phase change material films deserves further investigations.

6.5 Perspectives

As discussed in the present chapter, the performances of PCM devices are intimately related to the physical and crystallization properties of phase change materials. In that context, a huge research effort has been devoted to understanding the phase transformations and to identify new phase change materials with optimized properties. Continuous progresses in characterization techniques and computer simulations allowed to get a better insight on the structure of the amorphous and crystalline phases, on the origin of the resistivity contrast between amorphous and crystalline states and on the crystallization mechanisms. The current trend for PCMs is to confine, and also reduce, the active volume. This requires to master interface effects and in particular to suppress or at least control oxidation effects that deeply affect the crystallization mechanisms. Besides, the combination of a high crystallization speed during programming and a high thermal stability of the amorphous phase at the operating temperature is required in PCMs. In that context, studies on crystallization kinetics and on the fragility of the amorphous phase still deserve further

investigations. Last but not least, a major issue for PCM technology is the aging of the amorphous phase and the related so-called resistance drift phenomenon. In literature, the drift is often attributed to a "structural relaxation," but its nature, despite recent efforts, is still unclear. Solving this issue is required before aiming at controlling or ideally reducing drift, in particular in small confined structures.

Acknowledgments We thank Dr. J.-Y. Raty (FNRS, University of Liège) for his constant and friendly support and numerous fruitful discussions on phase change material science. P. N. acknowledges the PCM team in LETI. This work has been supported in part by ST-LETI bilateral program A, European 621217-PANACHE project, LabEx MINOS ANR-10-LABX-55-01, and the SESAME project ANR-15-CE24-0021. SOLEIL, ESRF, and the Italian beamline LISA at ESRF are acknowledged for granting synchrotron beam time. Dr. F. D'Acapito (CNR-IOM-OGG c/o ESRF) is also acknowledged for his support during EXAFS experiments, friendship, and outstanding XAS data analysis expertise. We sincerely acknowledge the past and present students and our long-standing close collaborators and also friends who contributed to this work: G.E. Ghezzi, M. Aoukar, M. Marra, P.-D. Szkutnik, P. Kowalczyk, P. Montéléon, C. Sabbione, M. Bernard, L. Fellouh, F. Fillot, N. Bernier, C. Licitra, V. Sousa, G. Navarro, V. Jousseaume, P.-H. Haumesser, and all those we may have forgotten.

References

1. J. Akola, R.O. Jones, Phys. Rev. B **76**, 235201 (2007)
2. J. Akola, R.O. Jones, Phys. Rev. B **79**, 134118 (2009)
3. J. Akola, J. Larrucea, R.O. Jones, Phys. Rev. B **83**, 094113 (2011)
4. J. Akola, R.O. Jones, Phys. Status Solidi B **249**, 1851 (2012)
5. C.A. Angell, Science **267**, 1924 (1995)
6. C.A. Angell, C.T. Moynihan, M. Hemmati, J. Non-Cryst. Solids **274**, 319 (2000)
7. S.K. Bahl, K.L. Chopra, J. Appl. Phys. **40**, 4940 (1969)
8. S.K. Bahl, K.L. Chopra, J. Appl. Phys. **41**, 2196 (1970)
9. G. Baldi, A. Fontana, G. Monaco, L. Orsingher, S. Rols, F. Rossi, B. Ruta, Phys. Rev. Lett. **102**, 195502 (2009)
10. G.B. Beneventi, L. Perniola, A. Fantini, D. Blachier, A. Toffoli, E. Gourvest, S. Maitrejean, V. Sousa, C. Jahan, J.F. Nodin, A. Persico, S. Loubriat, A. Roule, S. Lhostis, H. Feldis, G. Reimbold, T. Billon, B. De Salvo, L. Larcher, P. Pavan, D. Bensahel, P. Mazoyer, R. Annunziata, F. Boulanger, Proc. Eur. Solid State Device Res. Conf. ESSDERC **2010**, 313 (2010)
11. G. Betti Beneventi, L. Perniola, V. Sousa, E. Gourvest, S. Maitrejean, J.C. Bastien, A. Bastard, B. Hyot, A. Fargeix, C. Jahan, J.F. Nodin, A. Persico, A. Fantini, D. Blachier, A. Toffoli, S. Loubriat, A. Roule, S. Lhostis, H. Feldis, G. Reimbold, T. Billon, B. De Salvo, L. Larcher, P. Pavan, D. Bensahel, P. Mazoyer, R. Annunziata, P. Zuliani, F. Boulanger, Solid State Electron. **197**, 65–66 (2011)
12. V. Bragaglia, F. Arciprete, W. Zhang, A.M. Mio, E. Zallo, K. Perumal, A. Giussani, S. Cecchi, J.E. Boschker, H. Riechert, S. Privitera, E. Rimini, R. Mazzarello, R. Calarco, Sci Rep **6**, 23843 (2015)
13. M. J. Breitwisch, in *Phase Change Materials: Science and Applications*, ed. by S. Raoux, M. Wuttig (Springer, New York, 2009), pp. 381–408
14. G. Bruns, P. Merkelbach, C. Schlockermann, M. Salinga, M. Wuttig, T.D. Happ, J.B. Philipp, M. Kund, Appl. Phys. Lett. **95**, 043108 (2009)
15. G.W. Burr, M.J. Breitwisch, M. Franceschini, D. Garetto, K. Gopalakrishnan, B. Jackson, B. Kurdi, C. Lam, L.a. Lastras, A. Padilla, B. Rajendran, S. Raoux, R.S. Shenoy, J. Vac. Sci. Technol. B: Microelectron. Nanom. Struct. **28**, 223 (2010)

16. M.A. Caldwell, S. Raoux, R.Y. Wang, H.-S.P. Wong, D.J. Milliron, J. Mater. Chem. **20**, 1285 (2010)
17. S. Caravati, M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello, Appl. Phys. Lett. **91**, 171906 (2007)
18. E. Carria, A.M. Mio, S. Gibilisco, M. Miritello, F. d'Acapito, M.G. Grimaldi, E. Rimini, Electrochem. Solid-State Lett. **14**, H480 (2011)
19. S. Caravati, D. Colleoni, R. Mazzarello, T.D. Kühne, M. Krack, M. Bernasconi, M. Parrinello, J. Phys. Condens. Matter **23**, 265801 (2011)
20. T. Chatterji, C.M.N. Kumar, U.D. Wdowik, Phys. Rev. B **91**, 054110 (2015)
21. T. Chattopadhyay, J.X. Boucherle, H.G.v. Schnering, J. Phys. C Solid State Phys. **20**, 1431 (1987)
22. M. Chen, K.A. Rubin, R.W. Barton, Appl. Phys. Lett. **49**, 502 (1986)
23. H.Y. Cheng, T.H. Hsu, S. Raoux, J.Y. Wu, P.Y. Du, M. Breitwisch, Y. Zhu, E.K. Lai, E. Joseph, S. Mittal, R. Cheek, a. Schrott, S.C. Lai, H.L. Lung, C. Lam, IEEE Int. Electron Devices Meet. 2011, 3. **4**, 1 (2011)
24. A.I. Chumakov, G. Monaco, A. Monaco, W.A. Crichton, A. Bosak, R. Rüffer, A. Meyer, F. Kargl, L. Comez, D. Fioretto, H. Giefers, S. Roitsch, G. Wortmann, M.H. Manghnani, A. Hushur, Q. Williams, J. Balogh, K. Parlinski, P. Jochym, P. Piekarz, Phys. Rev. Lett. **106**, 225501 (2011)
25. E. Cho, S. Han, D. Kim, H. Horii, H.S. Nam, J. Appl. Phys. **109**, 1 (2011)
26. J.L.F. Da Silva, A. Walsh, H. Lee, Phys. Rev. B **78**, 224111 (2008)
27. V.L. Deringer, W. Zhang, M. Lumeij, S. Maintz, M. Wuttig, R. Mazzarello, R. Dronskowski, Angew. Chem. Int. Ed. **53**, 10817 (2014)
28. V.L. Deringer, R. Dronskowski, M. Wuttig, Adv. Funct. Mater. **25**, 6343 (2015)
29. K. Do, D. Lee, D.H. Ko, H. Sohn, M.H. Cho, Electrochem. Solid-State Lett. **13**, H284 (2010)
30. A.H. Edwards, A.C. Pineda, P.a. Schultz, M.G. Martin, A.P. Thompson, H.P. Hjalmarson, C.J. Umrigar, Phys. Rev. B **73**, 045210 (2006)
31. T. Egami , S.J.L Billinge, *Underneath the Bragg peaks, Structural Analysis of Complex Materials*, ed. By R.W. Cahn (Elsevier, Pergamon Materials Series Oxford, 2003)
32. J.M. van Eijk, *Structural Analysis of Phase-Change Materials Using X-Ray Absorption Measurements* (RWTH Aachen University, PhD thesis, 2010.) http://publications.rwth-aachen.de/record/64226/files/3595.pdf
33. R. Fallica, E. Varesi, L. Fumagalli, S. Spadoni, M. Longo, C. Wiemer, Phys. Status Solidi - Rapid. Res. Lett. **7**, 1107 (2013)
34. A. Fantini, L. Perniola, M. Armand, J.F. Nodin, V. Sousa, A. Persico, J. Cluzel, C. Jahan, S. Maitrejean, S. Lhostis, A. Roule, C. Dressler, G. Reimbold, B. DeSalvo, P. Mazoyer, D. Bensahel, F. Boulanger, Proc. Int. Memory Workshop **2009**, 66–67 (2009)
35. A. Fantini, V. Sousa, L. Perniola, E. Gourvest, J.C. Bastien, S. Maitrejean, S. Braga, N. Pashkov, A. Bastard, B. Hyot, A. Roule, A. Persico, H. Feldis, C. Jahan, J.F. Nodin, D. Blachier, A. Toffoli, G. Reimbold, F. Fillot, F. Pierre, R. Annunziata, D. Benshael, P. Mazoyer, C. Vallée, T. Billon, J. Hazart, B. De Salvo, F. Boulanger, Tech. Dig. - Int. Electron devices meet. IEDM **104**, 644 (2010)
36. P. Fantini, M. Ferro, A. Calderoni, S. Brazzelli, Appl. Phys. Lett. **100**, 213506 (2012)
37. M.H. Flores-Ruiz, G.G. Naumis, J. Chem. Phys. **131**, 154501 (2009)
38. P. Fons, T. Matsunaga, A.V. Kolobov, M. Krbal, J. Tominaga, N. Yamada, Phys. Status Solidi B **249**, 1919 (2012)
39. I. Friedrich, V. Weidenhof, W. Njoroge, P. Franz, M. Wuttig, J. Appl. Phys. **87**, 4130 (2000)
40. S. Gabardi, S. Caravati, G.C. Sosso, J. Behler, M. Bernasconi, Phys. Rev. B **92**, 054201 (2015)
41. J.P. Gaspard, A. Pellegatti, F. Marinelli, C. Bichara, Philos. Mag. B **77**, 727 (1998)
42. J.P. Gaspard, Comptes Rendus Physique **17**, 389 (2016)
43. G.E. Ghezzi, J.Y. Raty, S. Maitrejean, A. Roule, E. Elkaim, F. Hippert, Appl. Phys. Lett. **99**, 151906 (2011)

44. G.E. Ghezzi, R. Morel, A. Brenac, N. Boudet, M. Audier, F. Fillot, S. Maitrejean, F. Hippert, Appl. Phys. Lett. **101**, 233113 (2012)
45. G.E. Ghezzi, P. Noé, M. Marra, C. Sabbione, F. Fillot, N. Bernier, J. Ferrand, S. Maîtrejean, F. Hippert, Appl. Phys. Lett. **104**, 221605 (2014)
46. J. González-Hernández, E. Prokhorov, Y. Vorobiev, J. Vac. Sci. Technol. A **18**, 1694 (2000)
47. E. Gourvest, B. Pelissier, C. Vallée, A. Roule, S. Lhostis, S. Maitrejean, J. Electrochem. Soc. **159**, H373 (2012)
48. J. Hegedüs, S.R. Elliott, Nat. Mater. **7**, 399 (2008)
49. Q. Hubert, C. Jahan, V. Sousa, L. Perniola, A. Kusiak, J.L. Battaglia, P. Noé, M. Bernard, M. Tessaire, F. Pierre, P. Zuliani, R. Annunziata, G. Pananakakis, B. De Salvo, International Conference on Solid State Devices and Materials **306**, 550–551 (2013)
50. Q. Hubert, C. Jahan, A. Toffoli, G. Navarro, S. Chandrashekar, P. Noé, D. Blachier, V. Sousa, L. Perniola, J.F. Nodin, A. Persico, R. Kies, S. Maitrejean, A. Roule, E. Henaff, M. Tessaire, P. Zuliani, R. Annunziata, G. Pananakakis, G. Reimbold, and B. De Salvo, 4th IEEE Int. Mem. Work. IMW 2012, pp. 1–4 (2012)
51. D. Ielmini, Y. Zhang, J. Appl. Phys. **102**, 054517 (2007)
52. D. Ielmini, S. Lavizzari, D. Sharma, A.L. Lacaita, Appl. Phys. Lett. **92**, 193511 (2008)
53. D. Ielmini, M. Boniardi, A.L. Lacaita, A. Redaelli, A. Pirovano, Microelectron. Eng. **86**, 1942 (2009)
54. H. Iwasaki, Y. Ide, M. Harigaya, Y. Kageyama, I. Fujimura, Jpn, J. Appl. Phys. **31**., Part 1, 2B, 461–465 (1992)
55. R. Jeyasingh, S.W. Fong, J. Lee, Z. Li, K.W. Chang, D. Mantegazza, M. Asheghi, K.E. Goodson, H.S.P. Wong, Nano Lett. **14**, 3419 (2014)
56. Global Standards for the Microelectronics Industry (2004).[Online]. Available: http://www.jedec.org
57. T.H. Jeong, M.R. Kim, H. Seo, S.J. Kim, S.Y. Kim, J. Appl. Phys. **86**, 774 (1999)
58. T.H. Jeong, M.R. Kim, H. Seo, J.W. Park, C. Yeon, J. Japanese, Appl. Physics **39**, 2775 (2000)
59. P. Jóvári, I. Kaban, J. Steiner, B. Beuneu, A. Schöps, M.A. Webb, Phys. Rev. B **77**, 035202 (2008)
60. M.-C. Jung, Y.M. Lee, H.-D. Kim, M.G. Kim, H.J. Shin, K.H. Kim, S.A. Song, H.S. Jeong, C.H. Cho, M. Han, Appl. Phys. Lett. **91**, 083514 (2007)
61. J.A. Kalb, F. Spaepen, M. Wuttig, Appl. Phys. Lett. **84**, 5240 (2004)
62. J.A. Kalb, C.Y. Wen, F. Spaepen, H. Dieker, M. Wuttig, J. Appl. Phys. **98**, 054902 (2005)
63. J. Kalikka, J. Akola, R.O. Jones, Phys. Rev. B **90**, 184109 (2014)
64. M.J. Kang, T.J. Park, Y.W. Kwon, D.H. Ahn, Y.S. Kang, H. Jeong, S.J. Ahn, Y.J. Song, B.C. Kim, S.W. Nam, H.K. Kang, G.T. Jeong, C.H. Chung, Tech. Dig. - Int. Electron devices meet. IEDM **105**(21), 39 (2011)
65. V.G. Karpov, Y.A. Kryukov, I.V. Karpov, M. Mitra, Phys. Rev. B **78**, 052201 (2008)
66. T. Kato, K. Tanaka, Jpn, J. Appl. Phys. **44**, 7340 (2005)
67. K.F. Kelton, A.L. Greer, *Nucleation in Condensed Matter: Applications in Materials and Biology* (Elsevier, Pergamon Materials Series, Oxford, UK, 2010)
68. J.-H. Kim, M.R. Kim, Jpn. J. Appl. Phys. **37**(Part 1), 2116 (1998)
69. K. Kim, J.-C. Park, J.-G. Chung, S.A. Song, M.-C. Jung, Y.M. Lee, H.-J. Shin, B. Kuh, Y. Ha, J.-S. Noh, Appl. Phys. Lett. **89**, 243520 (2006)
70. I.S. Kim, S.L. Cho, D.H. Im, E.H. Cho, D.H. Kim, G.H. Oh, D.H. Ahn, S.O. Park, S.W. Nam, J.T. Moon, and C.H. Chung, Dig. Tech. Pap., 2010 Symp. VLSI Technol., 203 (2010)
71. K.H. Kim, S.J. Choi, J.G. Chung, J.H. Lee, S. Heo, Jpn. J. Appl. Phys. **49**, 0618011 (2010)
72. S. Kohara, K. Kato, S. Kimura, H. Tanaka, T. Usuki, H. Suzuya, H. Tanaka, Y. Moritomo, T. Matsunaga, N. Yamada, Y. Tanaka, H. Suematsu, M. Takata, Appl. Phys. Lett. **89**, 201910 (2006)
73. A.V. Kolobov, J. Tominaga, P. Fons, T. Uruga, Appl. Phys. Lett. **82**, 382 (2003)
74. A. Kolobov, P. Fons, A. Frenkel, A. Ankudinov, J. Tominaga, T. Uruga, Nat. Mater. **3**, 703 (2004)

75. A.V. Kolobov, P. Fons, J. Tominaga, A.L. Ankudinov, S.N. Yannopoulos, K.S. Andrikopoulos, J. Phys. Condens. Matter **16**, S5103 (2004)
76. A.V. Kolobov, P. Fons, J. Tominaga, A.I. Frenkel, A.L. Ankudinov, S.N. Yannopoulos, K.S. Andrikopoulos, T. Uruga, Jpn. J. Appl. Phys. **44**, 3345 (2005)
77. A.V. Kolobov, J. Tominaga, *Chalcogenides Metastability and Phase Change Phenomena*, vol 164 (Springer series In Materials Science, Berlin, 2012)
78. A. V. Kolobov (ed.), Phase-change memory: Science and Applications special issue. Phys. Status Solidi B **249**(10), 1817–2048 (2012)
79. A.V. Kolobov, P. Fons, B. Hyot, B. André, J. Tominaga, Y. Tamenori, H. Yoshikawa, K. Kobayashi, Appl. Phys. Lett. **100**, 061910 (2012)
80. A.V. Kolobov, P. Fons, J. Tominaga, Phys. Rev. B **87**, 155204 (2013)
81. B.J. Kooi, J.T.M. De Hosson, J. Appl. Phys. **92**, 3584 (2002)
82. U. Köster, Mater. Sci. Eng. **97**, 233 (1988)
83. M. Krbal, A.V. Kolobov, P. Fons, J. Tominaga, S.R. Elliott, J. Hegedus, T. Uruga, Phys. Rev. B **83**, 054203 (2011)
84. M. Krbal, A.V. Kolobov, P. Fons, K.V. Mitrofanov, Y. Tamenori, J. Hegedüs, S.R. Elliott, J. Tominaga, Appl. Phys. Lett. **102**, 111904 (2013)
85. A. Kusiak, J.-L. Battaglia, P. Noé, V. Sousa, F. Fillot, J. Phys. Conf. Ser. **745**, 032104 (2016)
86. P. La Fata, F. Torrisi, S. Lombardo, G. Nicotra, R. Puglisi, E. Rimini, J. Appl. Phys. **105**, 083546 (2009)
87. B.S. Lee, J.R. Abelson, S.G. Bishop, D.H. Kang, B.K. Cheong, K.B. Kim, J. Appl. Phys. **97**, 093509 (2005)
88. B.-S. Lee, S. G. Bishop, in Phase Change Materials – Science and Applications, ed. by S. Raoux and M. Wuttig (Springer US, 2009), pp. 175–198. DOI: 10.1007/978-0-387-84874-7
89. M. Lee, H.J. Shin, S.J. Choi, J.H. Oh, H.S. Jeong, K. Kim, M.-C. Jung, Curr. Appl. Phys. **11**, 710 (2011)
90. T.H. Lee, D. Loke, S.R. Elliott, Adv. Mater. **27**, 5477–5483 (2015)
91. D. Lencer, M. Salinga, B. Grabowski, T. Hickel, J. Neugebauer, M. Wuttig, Nat. Mater. **7**, 972 (2008)
92. D. Lencer, M. Salinga, M. Wuttig, Adv. Mater. **23**, 2030 (2011)
93. J.M. Li, H.X. Yang, K.G. Lim, Appl. Phys. Lett. **100**, 1 (2012)
94. J. Li, B. Luan and C. Lam, 2012 IEEE International Reliability Physics Symposium (IRPS), Anaheim, CA, (2012), pp. 6C.1.1–6C.1.6
95. R.G. Liang, C.B. Peng, K. Nagata, K. Daly-Flynn, M. Mansuripur, Appl. Opt. **41**, 370–378 (2002)
96. X.Y. Liu, J. Chem. Phys. **112**, 9949 (2000)
97. J. Luckas, A. Olk, P. Jost, H. Volker, J. Alvarez, A. Jaffré, P. Zalden, A. Piarristeguy, A. Pradel, C. Longeaud, M. Wuttig, Appl. Phys. Lett. **105**, 092108 (2014)
98. Y. Maeda, M. Wakagi, Jpn. J. Appl. Phys. **30**(1), 101–106 (1991)
99. Chapters 15–18 in Molecular Dynamics Simulations of Disordered Materials: From Network Glasses to Phase-Change Memory Alloys, Ed. By Carlo Massobrio, Jincheng Du, Marco Bernasconi, Philip S. Salmon (Springer Series in Materials Science 215, Springer International Publishing, Switzerland, 2015)
100. T. Matsunaga, N. Yamada, Phys. Rev. B **69**, 104111 (2004)
101. T. Matsunaga, N. Yamada, Y. Kubota, Acta Crystallogr. Sect. B Struct. Sci. **60**, 685 (2004)
102. T. Matsunaga, R. Kojima, N. Yamada, K. Kifune, Y. Kubota, Y. Tabata, M. Takata, Inorg. Chem. **45**, 2235 (2006)
103. T. Matsunaga, H. Morita, R. Kojima, N. Yamada, K. Kifune, Y. Kubota, Y. Tabata, J.J. Kim, M. Kobata, E. Ikenaga, K. Kobayashi, J. Appl. Phys. **103**, 093511 (2008)
104. R. Mazzarello, S. Caravati, S. Angioletti-Uberti, M. Bernasconi, M. Parrinello, Phys. Rev. Lett. **104**, 085503 (2010)
105. R.E. Meinders, A.V. Mijiritskii, L. van Pieterson, M. Wuttig, *Optical Data Storage – Phase-Change Media and Recording* (Springer, Netherlands, 2006)

106. X.S. Miao, T.C. Chong, Y.M. Huang, K.G. Lim, P.K. Tan, L.P. Shi, Jpn. J. Appl. Phys. **38**, 1638 (1999)
107. M. Micoulaut, J.Y. Raty, C. Otjacques, C. Bichara, Phys. Rev.B **81**, 174206 (2010)
108. M. Mitra, Y. Jung, D.S. Gianola, R. Agarwal, Appl. Phys. Lett. **96**, 222111 (2010)
109. K.V. Mitrofanov, A.V. Kolobov, P. Fons, X. Wang, J. Tominaga, Y. Tamenori, T. Uruga, N. Ciocchini, D. Ielmini, Appl. Phys. **115**, 173501 (2014)
110. L.J. Moore, D.R. MacFarlane, P.J. Newman, J. Non-Cryst. Solids **140**, 159 (1992)
111. S. Mukhopadhyay, J. Sun, A. Subedi, T. Siegrist, D.J. Singh, Sci Rep **6**, 25981 (2016)
112. G. Navarro, V. Sousa, A. Persico, N. Pashkov, A. Toffoli, J.C. Bastien, L. Perniola, S. Maitrejean, A. Roule, P. Zuliani, R. Annunziata, B. De Salvo, Solid State Electron. **89**, 93 (2013)
113. G.G. Naumis, H.M. Flores-Ruiz, Phys. Rev. B **78**, 094203 (2008)
114. W.K. Njoroge, H.-W. Wöltgens, M. Wuttig, J. Vac. Sci. Technol. A **20**(1), 230 (2002)
115. P. Noé, C. Sabbione, N. Castellani, G. Veux, G. Navarro, V. Sousa, F. Hippert, F. d'Acapito, J. Phys. D. Appl. Phys. **49**, 035305 (2016)
116. P. Noé, C. Sabbione, N. Bernier, N. Castellani, F. Fillot, F. Hippert, Acta Mater. **110**, 142 (2016)
117. T. Nonaka, G. Ohbayashi, Y. Toriumi, Y. Mori, H. Hashimoto, Thin Solid Films **370**, 258 (2000)
118. K. Ohara, L. Temleitner, K. Sugimoto, S. Kohara, T. Matsunaga, L. Pusztai, M. Itou, H. Ohsumi, R. Kojima, N. Yamada, T. Usuki, A. Fujiwara, M. Takata, Adv. Funct. Mater. **22**, 2251 (2012)
119. N. Ohshima, J. Appl. Phys. **79**, 8357 (1996)
120. J.K. Olson, H. Li, T. Ju, J.M. Viner, P.C. Taylor, J. Appl. Phys. **99**, 103508 (2006)
121. J. Orava, T. Wagner, J. Sik, J. Prikryl, M. Frumar, L. Benes, J. Appl. Phys. **104**, 43523 (2008)
122. J. Orava, a.L. Greer, B. Gholipour, D.W. Hewak, C.E. Smith, Nat. Mater. **11**, 279 (2012)
123. J. Orava, D.W. Hewak, A.L. Greer, Adv. Funct. Mater. **25**, 4851 (2015)
124. M.A. Otooni, *Science and Technology of Rapid Solidification and Processing* (NATO Science Series E vol. 278, Springer, Netherlands, (2012)
125. S.R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968)
126. J.-W. Park, S.H. Eom, H. Lee, J.L.F. Da Silva, Y.-S. Kang, T.-Y. Lee, Y.H. Khang, Phys. Rev. B **80**, 115209 (2009)
127. L. Pauling, *Nature of Chemical Bond* (Cornell Univ. Press, New York, 1939)
128. C. Peng, L. Wu, F. Rao, Z. Song, X. Zhou, M. Zhu, B. Liu, D. Yao, S. Feng, P. Yang, J. Chu, Scr. Mater. **65**, 327 (2011)
129. L. Perniola, V. Sousa, A. Fantini, E. Arbaoui, A. Bastard, M. Armand, A. Fargeix, C. Jahan, J.F. Nodin, A. Persico, D. Blachier, A. Toffoli, S. Loubriat, E. Gourvest, G. Betti Beneventi, H. Feldis, S. Maitrejean, S. Lhostis, A. Roule, O. Cueto, G. Reimbold, L. Poupinet, T. Billon, B. De Salvo, D. Bensahel, P. Mazoyer, R. Annunziata, P. Zuliani, F. Boulanger, IEEE Electron Device Lett. **31**, 488 (2010)
130. A. Pirovano, A.L. Lacaita, F. Pellizzer, S.A. Kostylev, A. Benvenuti, R. Bez, IEEE Trans. on Electr. Devices **51**, 714 (2004)
131. S. Privitera, E. Rimini, R. Zonca, Appl. Phys. Lett. **85**, 3044 (2004)
132. M. Putero, M.-V. Coulet, C. Muller, C. Baethz, S. Raoux, H.-Y. Cheng, Appl. Phys. Lett. **108**, 101909 (2016)
133. S. Raoux, J.L. Jordan-Sweet, A.J. Kellock, J. Appl. Phys. **103**, 114310 (2008)
134. S. Raoux, C. Cabral Jr., L. Krusin-Elbaum, J.L. Jordan-Sweet, K. Virwani, M. Hitzbleck, M. Salinga, A. Madan, T.L. Pinto, J. Appl. Phys. **105**, 064918 (2009)
135. Phase Change Materials – Science and Applications, ed. by S. Raoux and M. Wuttig (Springer (2009)
136. S. Raoux, H.-Y. Cheng, B. Muñoz, J.L. Jordan-Sweet, Proc. Eur. Phase Chang. Ovonic Sci. Symp. **2009**, 91 (2009)
137. S. Raoux, H.Y. Cheng, M.A. Caldwell, H.S.P. Wong, Appl. Phys. Lett. **95**, 071910 (2009)

138. S. Raoux, A.K. Konig, H.Y. Cheng, D. Garbin, R.W. Cheek, J.L. Jordan-Sweet, M. Wuttig, Phys. Status Solidi B **249**, 1999 (2012)
139. J.-Y. Raty, C. Otjacques, J.-P. Gaspard, C. Bichara, Solid State Sci. **12**, 193 (2010)
140. J.Y. Raty, P. Noé, G. Ghezzi, S. Maîtrejean, C. Bichara, F. Hippert, Phys. Rev. B **88**, 014203 (2013)
141. J.Y. Raty, W. Zhang, J. Luckas, C. Chen, R. Mazzarello, C. Bichara, M. Wuttig, Nat. Commun. **6**, 1 (2015)
142. J.J. Rehr, Rev. Mod. Phys. **72**, 621 (2000)
143. J. Robertson, B. Huang, Phys. Status Solidi B **249**, 1867 (2012)
144. I. Ronneberger, W. Zhang, H. Eshet, R. Mazzarello, Adv. Funct. Mater. **25**, 6407 (2015)
145. M. Rizzi, A. Spessot, P. Fantini, D. Ielmini, Appl. Phys. Lett. **99**, 223513 (2011)
146. M. Salinga, E. Carria, A. Kaldenbach, M. Bornhöfft, J. Benke, J. Mayer, M. Wuttig, Nat. Commun. **4**, 2371 (2013)
147. A. Schlieper, Y. Feutelais, S.G. Fries, B. Legendre, R. Blachnik, Calphad **23**(1), 1 (1999)
148. A. Sebastian, M. Le Gallo, D. Krebs, Nat. Commun. **5**, 1 (2014)
149. R.M. Shelby, S. Raoux, J. Appl. Phys. **105**, 104902 (2009)
150. K. Shportko, S. Kremers, M. Woda, D. Lencer, J. Robertson, M. Wuttig, Nat. Mater. **7**, 653 (2008)
151. T. Siegrist, P. Jost, H. Volker, M. Woda, P. Merkelbach, C. Schlockermann, M. Wuttig, Nat. Mater. **10**, 202 (2011)
152. R.E. Simpson, M. Krbal, P. Fons, A.V. Kolobov, J. Tominaga, T. Uruga, H. Tanida, Nano Lett. **10**, 414419 (2010)
153. D.J. Singh, J. Appl. Phys. **113**, 203101 (2013)
154. G.C. Sosso, G. Miceli, S. Caravati, F. Giberti, J. Behler, M. Bernasconi, J. Phys. Chem. Lett. **4**, 4241 (2013)
155. G.C. Sosso, M. Salvalaglio, J. Behler, M. Bernasconi, M. Parrinello, J. Phys. Chem. C **119**, 6428 (2015)
156. V. Sousa, G. Navarro, N. Castellani, M. Coue, O. Cueto, C. Sabbione, P. Noe, L. Perniola, S. Blonkowski, P. Zuliani, R. Annunziata, Dig. Tech. Pap. - Symp. VLSI Technol. **2015**, T98 (2015)
157. J.R. Stellhorn, S. Hosokawa, W.-C. Pilgrim, N. Blanc, N. Boudet, H. Tajiri, S. Kohara, Phys. Status Solidi B **253**, 1038 (2016)
158. Z. Sun, J. Zhou, R. Ahuja, Phys. Rev. Lett. **96**, 1 (2006)
159. J. Tominaga, T. Nakano, N. Atoda, Jpn. J. Appl. Phys. **37**, 1852 (1998)
160. W. Wang, D. Loke, L. Shi, R. Zhao, H. Yang, L.-T. Law, L.-T. Ng, K.-G. Lim, Y.-C. Yeo, T.-C. Chong, A.L. Lacaita, Sci Rep **2**, 360 (2012)
161. X. Wei, L. Shi, T.C. Chong, R. Zhao, L.H. Koon, Jpn. J. Appl. Phys. **46**, 2211 (2007)
162. W. Wełnic, S. Botti, L. Reining, M. Wuttig, Phys. Rev. Lett. **98**, 236403 (2007)
163. M. Wimmer, M. Kaes, C. Dellen, M. Salinga, Front. Phys. **2**, 75–71 (2014)
164. M. Wuttig, N. Yamada, Nat. Mater. **6**, 824 (2007)
165. M. Wuttig, D. Lüsebrink, D. Wamwangi, W. Wełnic, M. Gillessen, R. Dronskowski, Nat. Mater. **6**, 122 (2007)
166. L. Xu, L. Tong, L. Geng, F. Yang, J. Xu, W. Su, D. Liu, Z. Ma, K. Chen, J. Appl. Phys. **110**, 013703 (2011)
167. N. Yamada, E. Ohno, N. Akahira, K. Nishiuchi, K. Nagat, M. Takao, Jpn. J. Appl. Phys. **26**. Supplement **61**, 26–24 (1987)
168. N. Yamada, E. Ohno, K. Nishiuchi, N. Akahira, M. Takao, J. Appl. Phys. **69**, 2849 (1991)
169. N. Yamada, T. Matsunaga, J. Appl. Phys. **88**, 7020 (2000)
170. N. Yamada, Phys. Status Solidi B **249**, 1837 (2012)
171. L.V. Yashina, R. Püttner, V.S. Neudachina, T.S. Zyubina, V.I. Shtanov, M.V. Poygin, J. Appl. Phys. **103**, 094909 (2008)
172. Y. Yin, H. Zhang, S. Hosaka, Y. Liu, Q. Yu, J. Phys. D. Appl. Phys. **46**, 505311 (2013)
173. P. Zalden, A. von Hoegen, P. Landreman, M. Wuttig, A.M. Lindenberg, Chem. Mater. **27**, 5641 (2015)
174. A.R. Zanatta, I. Chambouleyron, Phys. Rev. B **48**, 4560 (1993)

175. W. Zhang, A. Thiess, P. Zalden, R. Zeller, P.H. Dederichs, J.-Y. Raty, M. Wuttig, S. Blügel, R. Mazzarello, Nat. Mater. **11**, 952 (2012)
176. W. Zhou, L. Wu, X. Zhou, F. Rao, Z. Song, D. Yao, W. Yin, S. Song, B. Liu, B. Qian, S. Feng, Appl. Phys. Lett. **105**, 243113 (2014)
177. X. Zhou, M. Xia, F. Rao, L. Wu, X. Li, Z. Song, S. Feng, H. Sun, ACS Appl. Mater. Interfaces **6**, 14207 (2014)
178. X. Zhou, W. Dong, H. Zhang, R.E. Simpson, Sci Rep **5**, 11150 (2015)
179. F. Zipoli, D. Krebs, A. Curioni, Phys. Rev. B **93**, 115201 (2016)
180. P. Zuliani, E. Varesi, E. Palumbo, M. Borghi, I. Tortorelli, D. Erbetta, G.D. Libera, N. Pessina, A. Gandolfo, C. Prelini, L. Ravazzi, R. Annunziata, IEEE Trans. Electron Devices **60**, 4020 (2013)

Chapter 7
Material Engineering for PCM Device Optimization

Véronique Sousa and Gabriele Navarro

7.1 Introduction

Phase change memories (PCMs) are among the most promising technologies for future generations of nonvolatile memories. The principal advantages of this technology are the low programming voltages, the very short reading and programming times, the good shrinking prospects, and the low manufacturing cost due to the reduced number of required masks levels. To date, the performances of several high-capacity demonstrators have been published: Numonyx [1] and Hynix [2] presented a 1 GB memory, while Samsung [3, 4] reported the results obtained with an 8 GB memory. Technology developments have also reached a high maturity level, thus enabling the first PCM product introduction as memory chips in cell phones [5, 6].

The operating fundamentals of a PCM device rely on the amorphous to crystalline phase transformation of a small volume of phase change material. PCMs owe their success to the unique combination of properties of the phase change materials, among which the large electrical contrast between the amorphous and the crystalline phase, the high crystallization speed of the amorphous phase, and the stability of the two programmed states at the user time scale. The PCMs concept was originally proposed in the 1960s by S. Ovshinsky [7], the first publication referring to the alloy $Ge_{10}Si_{12}As_{30}Te_{48}$ as a switching material. Later, in the 1980s, research works highlighted the ternary system GeSbTe (GST) with different compositions located on the pseudo-binary line $GeTe/Sb_2Te_3$, allowing crystallization times as short as a few tens of nanoseconds [8]. The stoichiometric compound $Ge_2Sb_2Te_5$ was initially developed for optical discs, taking advantage of the wide optical contrast between the ordered phases and disordered phase, before being used in

V. Sousa (✉) • G. Navarro
CEA-LETI, MINATEC Campus, Grenoble, France
e-mail: veronique.sousa@cea.fr

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_7

the late 1990s for PCM memory devices exploiting its large electrical contrast. On that date, this material had demonstrated an excellent endurance (over 10^{13} write/erase cycles) as well as the ability to store multiple levels of information in a memory [9]. However, PCM devices based on $Ge_2Sb_2Te_5$ are not able to achieve the thermal stability required for automotive applications (10 years at 150 °C) or for pre-encoding the data before the soldering process (2 min at 260 °C). Besides, high-performance applications require a very fast crystallization behavior, and a reduction in the programming current is desirable for the purpose of miniaturization of the cell. In the recent years, phase change material engineering focused in the GST system has enabled the optimization of the device performances in various directions, allowing to achieve either a reduction of the programming current, an increase of the programming speed, or an enhancement of the thermal stability of the programmed states. Few other investigations showed promising results in different material systems, among which the GaSb-based materials.

In this chapter, we present the material engineering studies performed so far, focusing on their contribution to the optimization of the PCM device performances. In Sect. 7.2, we first review the operating fundamentals of the PCM devices. This review allows us to highlight the phase change material characteristics, which control the PCM device performance and reliability. This analysis points out some conflicting material parameters, showing the necessary compromise to be found between, for example, the switching speed and the thermal stability of the programmed states. Then, we present state-of-the-art results from the material engineering investigations targeting the enhancement of the PCM device performances. After a brief review of the performances obtained with the reference compounds $Ge_2Sb_2Te_5$ and GeTe, we focus on the effect of additives which can be added to the stoichiometric reference alloys, thus allowing to achieve a reduction of the programming current, as well as an enhancement of the thermal stability of the programmed state. Then, we show how nonstoichiometric GST alloys can further be used to optimize the performances of the PCM devices. We first present the Ge-rich GST alloys which allow to enhance the thermal stability of the programed states and then the Sb-rich GST materials which are suitable for applications requiring a high programing speed. To finish, we review the device performances which have been obtained with materials outside the GST system, among which the GaSb-based materials.

7.2 Review of the Phase Change Material Properties Which Impact the PCM Device Performance and Reliability

The first property of the phase change materials required for the storage of a binary information in a PCM memory is the difference in electrical resistivity between the amorphous phase and the crystalline phase. Figure 7.1 shows the variation of the

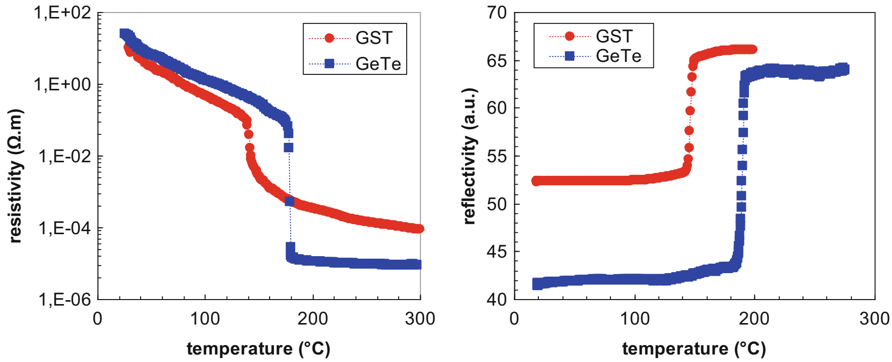


Fig. 7.1 Electrical resistivity and optical reflectivity variation as a function of the temperature for initially amorphous thin films of $Ge_2Sb_2Te_5$ and GeTe

electrical resistivity versus temperature for two typical phase change materials based on stoichiometric compounds: $Ge_2Sb_2Te_5$ and GeTe. The sharp drops in resistivity observed respectively at 150 °C and 180 °C correspond to the crystallization of the materials. In both cases, the resistivity difference between the amorphous and crystalline phases exceeds several orders of magnitude. Figure 7.1 also shows the variation of the optical reflectivity according to the temperature of the initially amorphous films of GeTe and $Ge_2Sb_2Te_5$, property used for reading the information stored on rewritable optical discs. Let us recall that according to Shportko et al. [10], the large difference between the optical and electrical characteristics of the amorphous and crystalline phases of several chalcogenide materials results from the fact that the conduction properties of the amorphous phase are those expected for a covalent semiconductor, while the conduction of the crystalline phase is greatly increased by a resonant bonding effect.

The second fundamental characteristic of a phase change material is the electronic switching of the amorphous phase. As already mentioned, in the PCM memory devices, the transition between the amorphous and the crystalline phase is induced by the Joule effect resulting from the application of a programming pulse to the storage element. The Joule heating of the amorphous phase is only possible, thanks to the electronic switching behavior observed when the amorphous material is submitted to a sufficiently high electric field. For a detailed description of this phenomenon and its physical origin, the reader is referred to Chap. 4.

Among the performances of PCM memory devices, the stability of the two programmed states in time and the speed of the writing process are of primary importance. These two characteristics are mostly controlled by the crystallization kinetics of the phase change material, and as a result, they appear as two conflicting device characteristics. This conflict is partly solved by the dependence in temperature of the crystallization speed, which helps to ensure on one side the stability of the amorphous phase at the low temperatures used for operating and storing the memory and on the other side a high speed of crystallization at the high temperatures used for programming. However, so far, dedicated materials have been

engineered to ensure either a high writing speed targeting, for example, high-performance applications, or a high thermal stability of the programmed states as required, for example, for embedded applications.

Crystallization times as small as a few nanoseconds have been demonstrated with phase change materials, e.g., GeTe [11]. Several interpretations have been proposed to account for these very short crystallization times [12, 13]. Note first that the phase change materials such as $Ge_2Sb_2Te_5$ or GeTe are stoichiometric compounds for which crystallization proceeds without any change in composition. In this case, the crystallization does not require long-distance diffusion of chemical species, which is a priori favorable to rapid crystallization process. Furthermore, the microstructure of the amorphous and crystalline phases plays a major role in the rate of crystallization, as shown by the publications based on experimental and theoretical approaches. Thus, rapid crystallization phenomenon appears to stem from the similarities between the short-range order (SRO) of the crystalline phase and the amorphous phase. Lencer [14] concluded from both atomistic calculations and experimental characterizations that the amorphous and crystalline phases have a similar energy and short-range order while having a very different density of states, thus explaining the optical and electrical contrast between the two phases.

As we already said, one major challenge that drives the development of PCM memories relates to improving the stability of the programmed states. For example, the specification in data retention for automotive applications is 10 years at 150 °C. Tougher still is the specification required in the case of pre-encoded memory components at the wafer level: in this case, the programmed information must be preserved during the soldering step of the component on the circuit board, which is for 2 min at 260 °C. Indeed, the thermal stability of the SET and RESET states can be compromised at high temperature by two physical phenomena:

- The ordering of the disordered phase material, that is to say the crystallization of the amorphous material [15]. This crystallization is accompanied by a decay of the resistance of the device and can lead to the loss of the high-resistance or RESET state.
- The structural relaxation in the disordered phase material, which is accompanied by an increase of the resistance referred to as “resistance drift” [16]. As expected, this phenomenon is obviously observed with the devices programmed in the RESET state in which the active material is amorphous and thus is highly disordered. But it is also observed on the devices in the SET state in which the material is polycrystalline. In this case, the drift of the resistance is interpreted as the result of the structural relaxation of some amorphous residues and attributed to the presence of grain boundaries [17]. Since it corresponds to a gradual increase in the resistance of the devices, this can lead to the loss of the SET state of low resistance, which will gradually get closer to the RESET state resistance level.

Once more, we will see that these two failure mechanisms unfortunately appear as competing phenomena: so far, experimental results have shown that materials allowing a higher stability of the RESET state also exhibit a higher drift of the SET

state [18]. The optimization of the thermal stability of the programmed states thus relates to the quest for a compromise.

Reducing the programming current, and more specifically the RESET current I_{rst}, is another leading research axis for phase change memories. As already explained in Chap. 9, the first solution to reduce the programming current is the miniaturization of the memory cell. Besides miniaturization, the reduction of the RESET current I_{rst} can be obtained by electrical and thermal optimization of the structure of the memory cells. Investigations have finally shown that with identical size and structure of the storage element, the programming current can be reduced through the optimization of the phase change material itself: lowering the melting temperature T_m decreases the power required to melt the material [19]; increasing the electrical resistivity of the crystalline phase ρ_c by composition modifications can increase the heat source [20]; SiO_2 additions in $Ge_2Sb_2Te_5$ are causing sharp reductions of the current I_{rst} [21] through the modification of the electrical and thermal properties of the material. Indeed, the programming current depends on a number of microscopic parameters characterizing the PCM material. We can identify these parameters through the review of the physical phenomena involved during programming.

One has to first consider the heat source at the origin of the temperature rise in the phase change material during the programming pulse.

Consider first the heat source at the origin of the temperature rise in the phase change material during the programming pulse, that is to say the Joule effect, which has been described in detail in Chap. 3. The power P_{PC} developed in the PCM layer is the sum of two contributions [22]: an ohmic part determined by the resistance R_{on} of the material in the ON conductive state and a non-ohmic part characterized by the holding voltage V_h:

$$P_{PC} = V_h \cdot I_{prog} + R_{on} \cdot I_{prog}{}^2 \tag{7.1}$$

Thus, a modification of the PCM material will affect the heat source during the programming pulse and thus the programming current through the modification of the resistance R_{on} of the material in the ON conducting state and the holding voltage V_h.

Next, consider the dissipation of heat during the RESET pulse, which can be assessed by the heat transfer equation [23]:

$$\rho C_P \cdot \frac{\partial T}{\partial t} + \mathrm{div}(-k \cdot \mathrm{grad}(T)) = j^2/\sigma \tag{7.2}$$

The dissipation of heat contains the change of energy due to the increase of the temperature of phase change material up to the melting temperature T_m, and the heat losses due to heat diffusion. Having reached the steady state during the first nanoseconds of the RESET pulse, a modification of the PCM material will affect the power required to melt the material through the modification of the melting temperature T_m and the thermal conductivity k_{th} [24].

Finally, it can also be important to consider the behavior of the material during the cooling period after the RESET pulse. In the case of a phase change material

having a high growth rate, it has been shown that a certain portion of PCM material located at the periphery of the molten volume could be recrystallized during cooling. The final amorphous volume is thus smaller than the melt volume, and overheating beyond the melting is necessary to obtain a high-resistance state. For example, the reduction of the current I_{prog}, following the reduction of the ratio Sb/Te in the PCM memory cells based on germanium-doped SbTe, is attributed to the decrease in the growth rate [25].

Aside from reducing the programming current and improving the performance in retention, other characteristics of the PCM devices can be improved by optimizing the phase change material. For example, it has been demonstrated that doping added to $Ge_2Sb_2Te_5$ can help enhancing the endurance. In fact, doping has been shown to retard the formation of holes (or "voids") and shift the positions of these holes to the center of the phase change material layer, thereby improving endurance by three orders of magnitude [26]. However, no details are given in the nature or amount of the doping in those investigations.

As far as material engineering is concerned, research works have also highlighted that the material fabrication method can be of outmost importance for taking full advantage of the material properties. For example, in the case of doped GST materials, Cheng and coworkers have pointed out the different device performances obtained when the doping element is added using co-sputtering from two targets or using a single composite target [27]. They highlight the necessity to control the doping process at the atomic level which allowed them to achieve at the same time on a 128 Mb test chip a high SET speed (20 ns), a high endurance (10^9 SET/RESET cycles), and a 65% RESET current reduction compared to un-doped GST. The nature of the doping element is again not mentioned.

In summary, material engineering can help to optimize several PCM device performances and reliability characteristics, among which the programming current, the programming speed, and the stability of the programmed states. The optimization work will also have to ensure that the overall operating characteristics of the devices remain suitable, including the resistance window, the endurance, etc. To date, theoretical arguments as well as experimental results show that compromise will have to be done between conflicting performances and reliability characteristics, thus orienting the choice of different classes of materials for different target applications.

7.3 Optimization of PCM Devices Integrating Stoichiometric Compounds from the GeSbTe System

7.3.1 The Reference Compounds $Ge_2Sb_2Te_5$ and GeTe

In the early 1990s, the company Matsushita selling optical discs under the Panasonic brand name raised interest in phase change materials with the

development of the alloy $Ge_2Sb_2Te_5$ which is still today the reference system for phase change materials. This alloy has been widely used for the rewritable optical disc of the DVD generation, allowing crystallization times of a few tens of nanoseconds and endurance of up to 10^{12} write/erase cycles. The $Ge_2Sb_2Te_5$ alloy corresponds to a stoichiometric compound, so that the crystallization of the alloy proceeds without any change in composition by the involvement of a nucleation process of seed crystals followed by the growth of these nuclei. It is the same for the Ge_1Te_1 or GeTe binary compound in the stoichiometry which is another widely studied material. These alloys fit therefore with the conditions of application of the Classical Nucleation Theory (CNT) frequently used to assess the nucleation and growth rates of a crystal within an amorphous matrix. The CNT can help us identifying the material characteristics which control the crystallization kinetics of the stoichiometric compounds and thus the crystallization speed and the thermal stability of the RESET state.

In the framework of the CNT, the nucleation and growth processes take place in a temperature range delimited by the glass transition temperature T_g, below which there is no atomic motion, and the melting temperature T_m, above which the crystalline phase is not thermodynamically stable. In this temperature range (T_g, T_m), the growth process is controlled by an energy barrier which is kinetic in nature since the crystallization is thermodynamically favorable. Nucleation is in turn controlled by both a kinetic and a thermodynamic energy barrier. The nucleation and growth rates exhibit both a maximum within the temperature range (T_g, T_m). The growth rate reaches its maximum value at a temperature close to the melting temperature, while the nucleation rate reaches its maximum value at a much lower temperature. The analysis of the expressions of the nucleation and growth rates given by the CNT highlights four parameters, which appear to play a major role in the crystallization process of stoichiometric compounds: T_g, T_m, and L_f and σ_{ac} [28]:

- The temperatures T_g and T_f define the temperature range wherein the nucleation and growth are possible. Kalb et al. [29] define the reduced glass transition temperature T_{rg} as the ratio of the glass transition temperature (approximately the crystallization temperature) to the melting temperature. They show that the value T_{rg} determines the crystallization mechanism: materials having a low value T_{rg} are called rapid nucleation materials since nucleation is effective throughout the heated amorphous volume, while the materials with higher value T_{rg} are called slow nucleating materials and crystallized by the motion of the amorphous/crystal interface.
- The energies L_f and σ_{ac} SET the maximum amplitude of the nucleation and growth rates. A larger value of the latent heat of fusion L_f corresponds to a greater stability of the crystalline phase with respect to the amorphous phase, thus resulting in a higher nucleation and growth rates. However, a larger value of L_f is usually accompanied by a higher interface energy σ_{ca}, which in fact reduces dramatically the amplitude of the nucleation rate. It is interesting to note the semi-empirical relationship proposed by Scapski and Turnbull [30, 31] which shows that the ratio σ_{ac}/L_f is almost constant.

Many authors have based the modeling of the crystallization of the phase change materials on the CNT. While acknowledging the usefulness of this approach which can reproduce the observed phenomena, it is worth noting here the limitations of this theory. Among the many assumptions that may be questioned, one of the strongest is certainly to consider the different microscopic parameters as constants, uniforms, and independent of the size of crystallites. A second limitation of this theory lies in the difficulty of assessing some of these parameters, like the interface energy between the amorphous and the crystalline phases σ_{ac} and the viscosity parameters.

As we have already mentioned, there are two emblematic stoichiometric compounds of phase change materials: $Ge_2Sb_2Te_5$ and GeTe. Figure 7.2 shows the variation of the nucleation and growth rates as a function of the temperature for the two reference compounds as calculated with the CNT. By applying this model, we get for the ternary compound $Ge_2Sb_2Te_5$ a high nucleation rate and for the binary compound GeTe a high growth rate, thus reproducing the experimental characteristics of these materials. In fact, through their crystallization characteristics, it is customary to refer to the $Ge_2Sb_2Te_5$ and GeTe compounds respectively as a nucleation-dominated material and as a growth-dominated material [32].

The crystallization curves of these two compounds as measured by the change in the square resistance of a thin film versus temperature (Fig. 7.1) reveal that the crystallization temperature of the stoichiometric alloy GeTe (180 °C) is higher than that of the $Ge_2Sb_2Te_5$ alloy (150 °C). However, the activation energy evaluated by the Kissinger method on thin films is slightly lower (Fig. 7.3).

The programming characteristics of PCM devices based on GeTe and $Ge_2Sb_2Te_5$ are shown in Fig. 7.4. Thanks to the very high growth rate of GeTe, devices based on this compound provide a very fast programming speed, presenting a resistance contrast over two orders of magnitude with SET programming pulses as short as 30 ns. The figure shows, for these two reference alloys, a similar programming current as well as a similar maximum temperature ensuring a retention of 10 years close to 100 °C. Thus, these alloys are not suitable for applications requiring a high thermal stability of the programmed states, like the automotive applications.

7.3.2 Light Element Additions in the Reference Compounds $Ge_2Sb_2Te_5$ and GeTe

Following the achievements of PCM devices based on stoichiometric compounds, research efforts have first turn to materials derived from the two reference alloys GeTe and $Ge_2Sb_2Te_5$. We report hereafter the performances of the devices obtained when adding light elements to the reference alloys, and the investigation performed with composite material formed by a mixture of the reference compounds with a dielectric material. The goal for modifying the stoichiometric compounds was

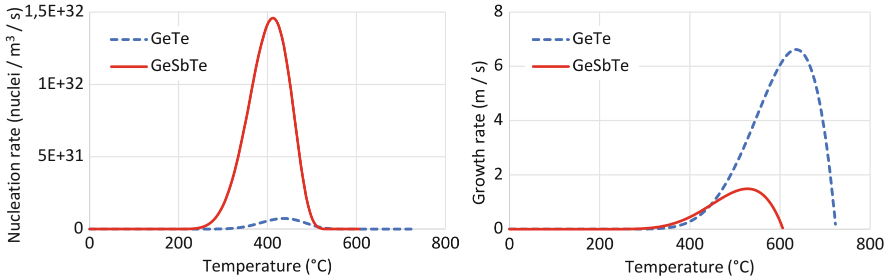


Fig. 7.2 Nucleation and growth rates in the stationary regime for the stoichiometric compounds $Ge_2Sb_2Te_5$ and GeTe as evaluated by the Classical Nucleation Theory

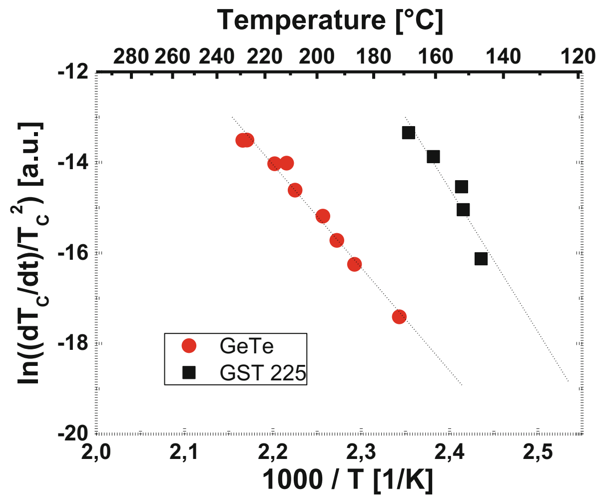


Fig. 7.3 Kissinger representation of the crystallization temperature as a function of the heating ramp, highlighting the higher activation energy for crystallization of GeTe versus Ge_2Sb_2Te

twofold: first, interfere with the crystallization process targeting a slowdown of its kinetics and a re-enforcement of the stability of the RESET state; second, increase the electrical and thermal resistivity of the PCM material to reduce the programming current via the enhancement of the heat source and the limitation of the heat dissipation.

7.3.2.1 Nitrogen

The addition of nitrogen has been experimented in several phase change materials. Nitrogen is easily introduced in the alloys by using reactive sputtering, the sputtering gas being a mixture of nitrogen and argon. By varying the ratio of the nitrogen and argon gas flux, the amount of nitrogen introduced in the deposited film can be tuned.

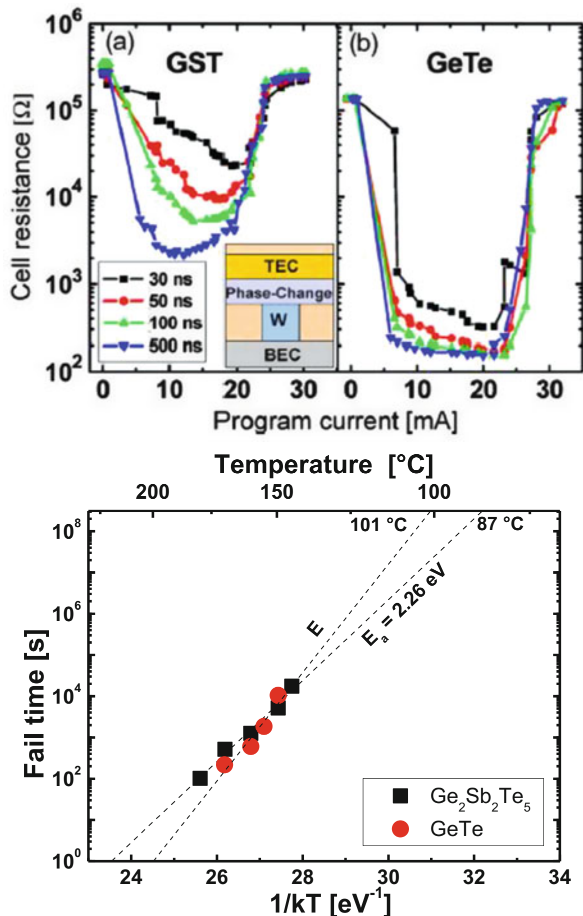


Fig. 7.4 Programming characteristics of PCM devices integrating the reference compounds $Ge_2Sb_2Te_5$ and GeTe (© 2016 IEEE. Reprinted, with permission, from [33]) and failure time of the RESET state versus temperature, highlighting the extrapolation of the data retention to 10 years

In the stoichiometric alloy GeTe, the addition of nitrogen generates an increase in the crystallization temperature as well as an increase of the activation energy of the crystallization (Fig. 7.5 and Table 7.1), thus illustrating the higher thermal stability of the amorphous phase of the material. Material characterization including the analysis of XPS spectra at the germanium and tellurium edges reveals that nitrogen preferentially binds with germanium atoms (see Chap. 7). The addition of nitrogen changes the X-ray diffraction spectra which are sensitive to the long-range order, indicating a transition from the rhombohedra phase to the cubic phase for nitrogen content exceeding 4%, and a decrease of the crystalline grain size. At the same time, the EXAFS spectra show that at short distance, the structure remains

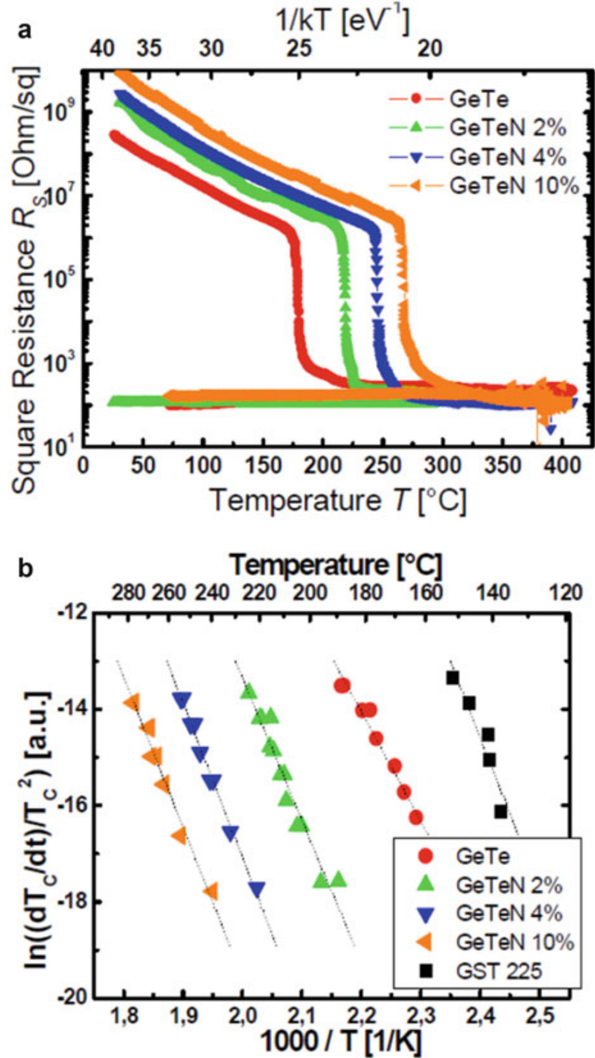


Fig. 7.5 Evaluation of the crystallization temperature and activation energy for crystallization for GeTe thin films, with and without nitrogen. (**a**) Square resistance of initially amorphous thin films as a function of the temperature (ramp of 10 °C/min). (**b**) Kissinger representation (© 2016 IEEE. Reprinted, with permission, from [35])

Table 7.1 Crystallization temperature T_c, melting temperature T_m (as evaluated by DSC measurements) and activation energy for crystallization E_A for several GeTeN alloys

Blanket material	T_c	T_m	E_A
(N%)	(°C)	(°C)	(eV)
0%	180	720	1.96 ± 0.102
2%	219		2.54 ± 0.184
4%	247	720	2.76 ± 0.095
10%	269	720	2.68 ± 0.150

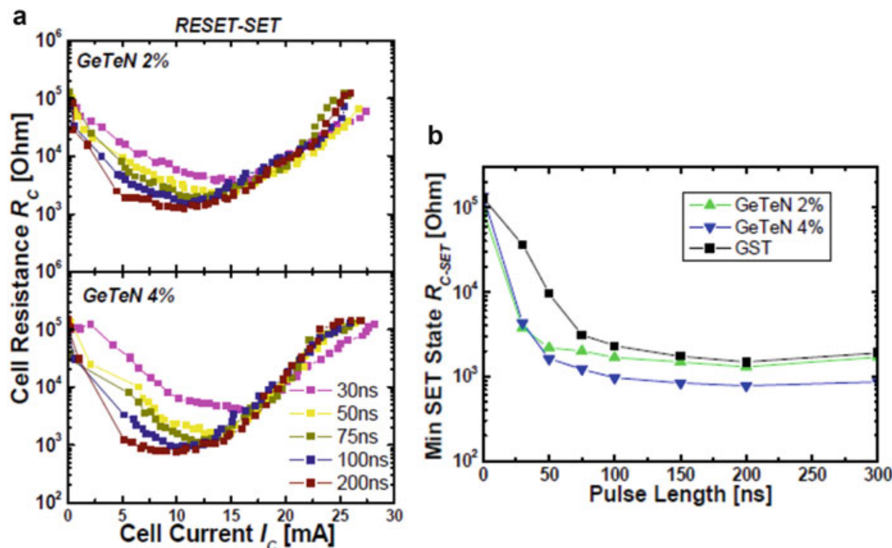


Fig. 7.6 (**a**) Programming characteristics R(i) of PCM devices integrating GeTeN alloys: each programming pulse is preceded by a RESET pulse of 30 mA and duration 100 ns. (**b**) Variation of the minimal resistance state as a function of the duration of the programming pulse (© 2016 IEEE. Reprinted, with permission, from [35])

locally rhombohedra. These results were interpreted by a bigger disorder at long distance accompanying the addition of nitrogen [34]. In addition, the introduction of nitrogen leads to the disappearance of the segregation of cubic germanium at high temperature, this resulting from the trapping of the germanium atoms in GeN bonds.

Regarding the performances of PCM devices integrating the alloy GeTeN, the programming current and the writing speed are shown to remain unchanged (Fig. 7.6). The programming characteristics RI show that starting from the RESET state, the successive transitions toward the SET and RESET states are not as steep as for devices based on the stoichiometric alloy GeTe, illustrating that nitrogen additions induces a strong nucleation process while the pure GeTe alloy is essentially dominated by growth. It seems that the initial composition of the thin films integrated in the devices is slightly enriched with germanium with respect to the stoichiometric composition $Ge_{50}Te_{50}$. Coombs [36] showed that while the stoichiometric alloy composition $Ge_{50}Te_{50}$ behaves as a nucleation-dominated material, nonstoichiometric GeTe alloys exhibit a crystallization process dominated by growth. Besides, the analysis performed on GeTeN thin film showed that nitrogen binds to the germanium. Thus the addition of nitrogen enables the GeTe material integrated in the devices to approach the stoichiometric composition $Ge_{50}Te_{50}$ by forming Ge-N bonds with the excess germanium atoms, hereafter favoring a crystallization behavior dominated by nucleation. As expected with the increase of the crystallization temperature and that of the activation energy of

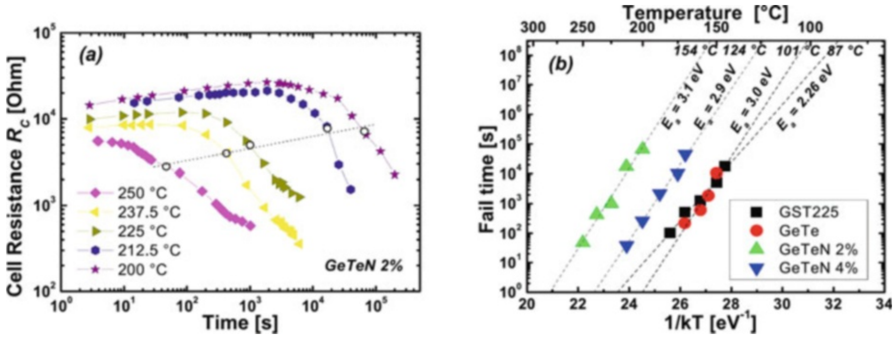


Fig. 7.7 Data retention for the RESET state of GeTeN devices. (**a**) Variation of the RESET resistance versus time. (**b**) Extrapolation to 10 years of the data retention in the Arrhenius representation

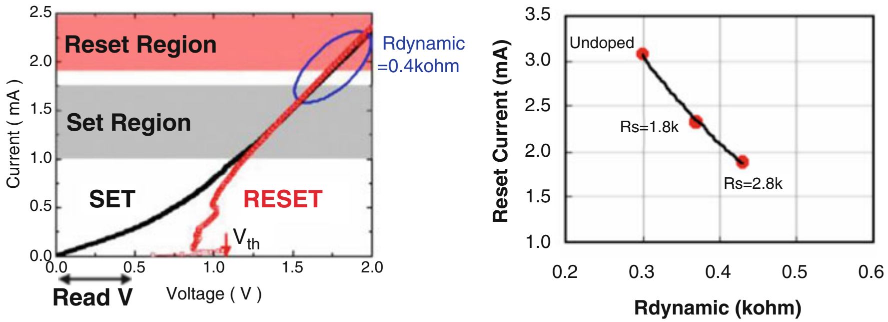


Fig. 7.8 Nitrogen doping into GST alloys increases the dynamic resistance of the PCM cell in the conductive ON state, resulting in lower writing current (© 2016 IEEE. Reprinted, with permission, from [39])

crystallization, the retention of the RESET state is largely enhanced with the addition of nitrogen (Fig. 7.7): the highest temperature for a 10-year data retention is extrapolated to be above 150 °C if 2% of nitrogen is added to GeTe.

Nitrogen doping was also experienced in $Ge_2Sb_2Te_5$-based devices. At the film level, an increase in the crystallization temperature is mentioned in the literature [37]. However, no report of thermal stability enhancement of the RESET state has been demonstrated at the device level. On the other side, nitrogen additions into $Ge_2Sb_2Te_5$ appear beneficial for the reduction of the RESET current (Fig. 7.8). The material was further successfully integrated in a 256 Mb demonstrator [38].

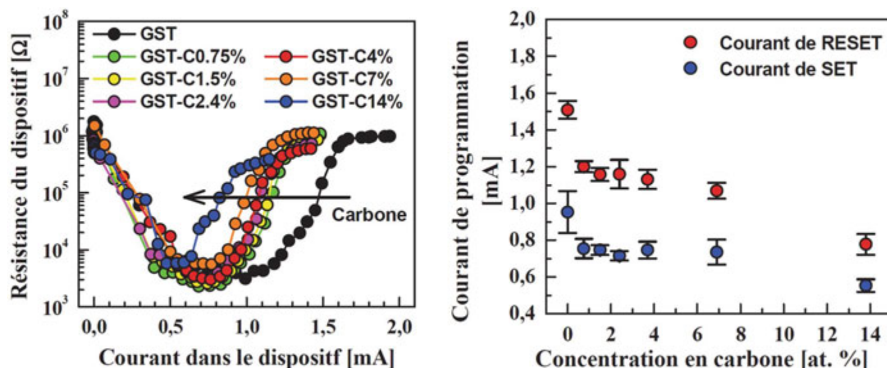


Fig. 7.9 Programming characteristics of PCM devices integrating GSTC alloys and variation of the RESET current as a function of the carbon content (Reprinted from the dissertation of Q. Hubert [41])

7.3.2.2 Carbon

In this section, we present the results obtained with the addition of carbon in the $Ge_2Sb_2Te_5$ stoichiometric compound. The material was deposited by co-sputtering using two targets: one of $Ge_2Sb_2Te_5$ and the other of carbon. The carbon content was modulated by varying the power applied to the targets. Carbon addition in the reference alloy $Ge_2Sb_2Te_5$ has first been shown to result in a reduction of the programming current [40]. As can be seen with the programming curves presented in Fig. 7.9, the addition of 5% carbon in the $Ge_2Sb_2Te_5$ alloy results in a significant reduction of more than 50% of the programming current. This current reduction is accompanied by a reduction of about 25% of the power delivered in the phase change material P_{PC} (Eq. 7.1).

To explain the reduction in programming power, we must consider either an increase of the power generated by the Joule effect through an increase of the dynamic resistance of the cell or a better thermal efficiency of the cell through a reduction of the thermal conductivity of the phase change material. The thermal conductivity measurements by modulated photothermal radiometry show that the addition of carbon reduces the thermal conductivity of the $Ge_2Sb_2Te_5$ alloy (Fig. 7.10). The measurement of the IV characteristics of PCM memory cells based on carbon-enriched $Ge_2Sb_2Te_5$ demonstrates that the addition of carbon also leads to an increase of the resistance R_{ON} of the PCM cells (Fig. 7.10). The consideration of both the lower thermal conductivity and the higher dynamic resistance in the simulations enables to reproduce the experimental programming curves (Fig. 7.11).

Thus, the addition of carbon into the alloy $Ge_2Sb_2Te_5$ simultaneously causes the decrease of the thermal conductivity of the material and the increase of its dynamic resistance. The reduction of the programming current results both from the reduction of the power required to reach the programming conditions and the increase of the heat source generated by the Joule effect. These property changes of the alloy

Température	$\kappa_{GST}/\kappa_{GST\text{-}C4\%}$
85 °C	2.4
135 °C	2.3
185 °C	2.0
265 °C	3.0
285 °C	3.0

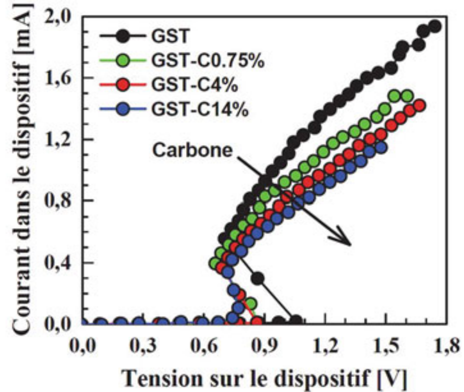


Fig. 7.10 Result of the evaluation of the thermal conductivity by modulated photothermal radiometry on GSTC thin films and IV characteristics of GSTC-based devices highlighting the increase of the resistance R_{ON} of GSTC-based devices (Reprinted from the dissertation of Q. Hubert [41])

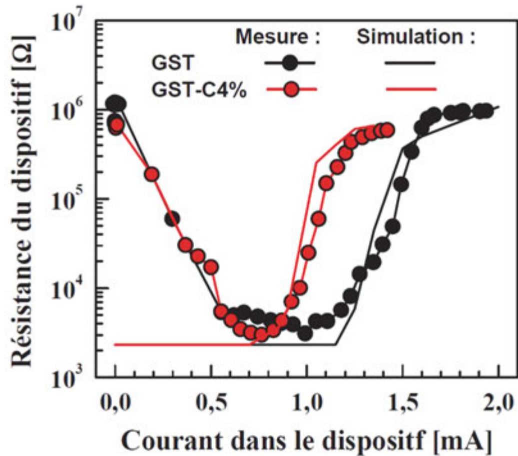


Fig. 7.11 Experimental and simulated programming characteristics of GSTC-based PCM devices. The simulation takes into account both the decrease of the thermal conductivity and the increase of the dynamic resistance of the devices (Reprinted from the dissertation of Q. Hubert [41])

$Ge_2Sb_2Te_5$ following the addition of carbon can be correlated to the reduction of the size of crystalline domains (see Chap. 7).

Apart from the current reduction, carbon additions in phase change materials have been shown to result in an enhancement of the thermal stability of the programmed states [42]. Figure 7.12 shows that the addition of carbon results in an increase in the crystallization temperature as measured on GeTe and $Ge_2Sb_2Te_5$ thin films. It is noted that the addition of about 25% of carbon allows the crystallization temperature of these alloys to approach the highest temperature that the wafers will see during the fabrication, which is 400 °C. In addition to a higher crystallization temperature, the better thermal stability of the amorphous phase is

also illustrated by the slowing down of the crystallization as illustrated by a less abrupt transition of the resistivity or the reflectivity at the crystallization.

Having demonstrated the greater thermal stability of the amorphous phase of the $Ge_2Sb_2Te_5$ compound including carbon additions, the improvement of the retention of the devices programmed in RESET state was verified (Fig. 7.13). The electrical results show that memory devices integrating GSTC alloys approach the specification required for automotive applications, which is 10 years at 150 °C. However, the thermal stability of GSTC alloys is still not sufficient to guarantee the stability of the information during the soldering step of the memory chip on the printed circuit board (2 min at 260 °C).

Further investigations revealed the combined effect of carbon and titanium in GeTe- and $Ge_2Sb_2Te_5$-based devices. Perniola and coworkers studied indeed the impact of a thin titanium layer inserted between the PCM layer and the TiN diffusion barrier [42]. This titanium layer is indeed necessary in the end product to optimize the adhesion of the upper electrode of TiN on the phase change material, and it has a significant impact on the crystallization of carbon-doped

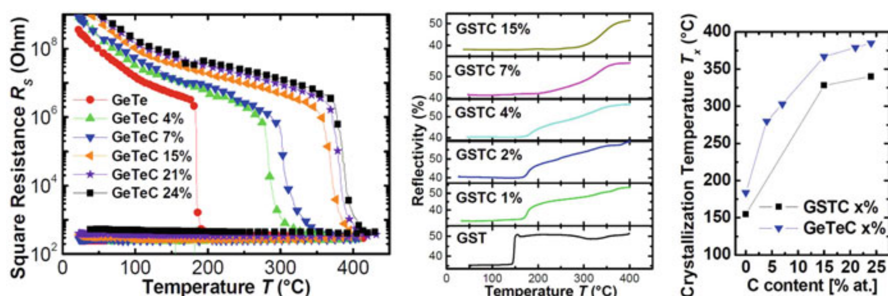


Fig. 7.12 Effect of the addition of carbon on the crystallization curves of GeTe and $Ge_2Sb_2Te_5$ thin films and crystallization temperature as a function of the carbon content (© 2016 IEEE. Reprinted, with permission, from [42])

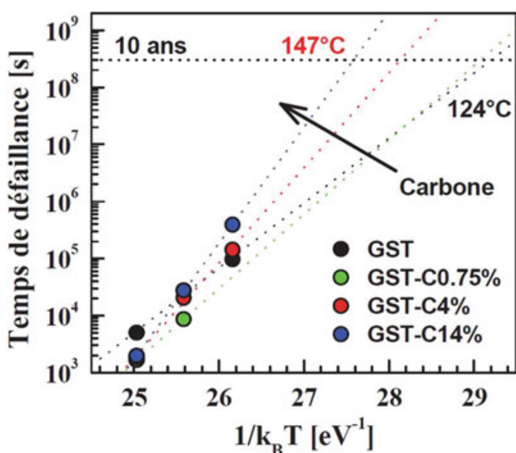


Fig. 7.13 Extrapolation at 10 years of the data retention for GSTC devices programmed in the RESET state (Reprinted from the dissertation of Q. Hubert [41])

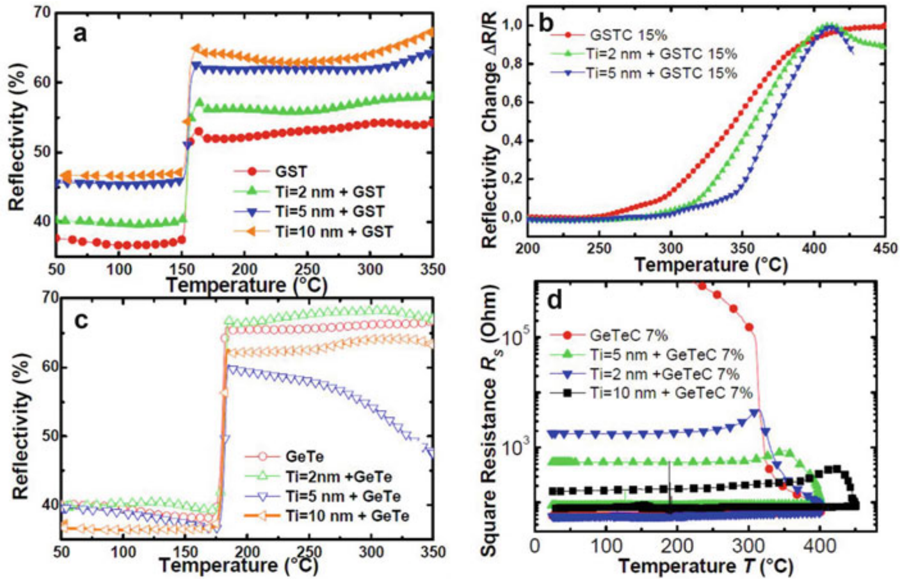


Fig. 7.14 Crystallization characteristics of GSTC and GeTeC thin films deposited on Ti sublayer, showing that the combination of carbon and titanium leads to an increase in crystallization temperature (© 2016 IEEE. Reprinted, with permission, from [42])

alloys. Figure 7.14 illustrates the impact of a titanium layer of variable thickness on the crystallization curves of GST and GeTe materials, doped or not with carbon. If the addition of the titanium layer does not impact the crystallization temperature of the un-doped alloys, there is on the contrary a significant increase of this temperature in the presence of carbon. For example, in the case of GeTe alloy doped with 7% of carbon, the crystallization temperature can grow by over 100 °C in the presence of a titanium layer of thickness 10 nm.

This increase in crystallization temperature was explained by the diffusion of titanium in the phase change material layer: the composition profiles, as measured by SIMS spectroscopy on a stack of GeTeC/Ti (Fig. 7.15) thin layers, reveal the interdiffusion of the different elements after annealing at 400 °C. On the composition profiles of the GSTC-based devices, integrating a thin layer of titanium and having been annealed at 400 °C during the final stages of manufacturing, the same behavior already identified with the thin films is observed, that is to say the diffusion of the titanium within the phase change material. They note here that the titanium diffusion is also accompanied by nonuniform profiles in germanium, tellurium, and carbon: the formation of Ti-Te bonds being favored [43], the diffusion of titanium generating a tellurium enrichment in the proximity of the upper electrode, and a germanium and carbon enrichment in the area close to the tungsten contact, that is to say in the active area of the device.

Following this material analysis, the thermal stability of the devices based on GSTC and incorporating a thin layer of titanium was investigated. Figure 7.16

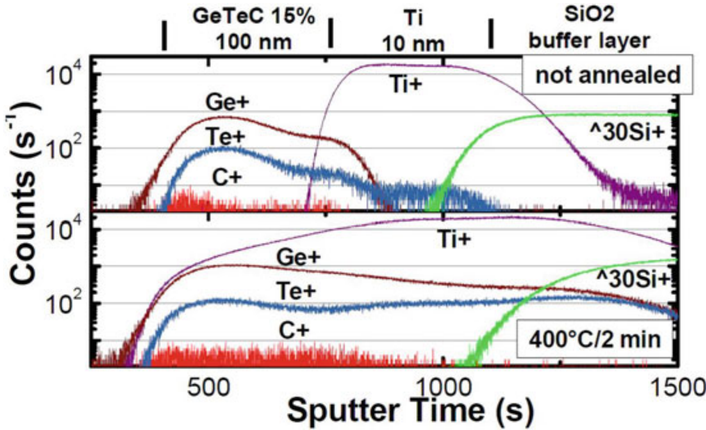


Fig. 7.15 SIMS profile measured in GeTeC 15% thin films deposited on a titanium layer, showing the interdiffusion of the layers after a 400 °C anneal (© 2016 IEEE. Reprinted, with permission, from [42])

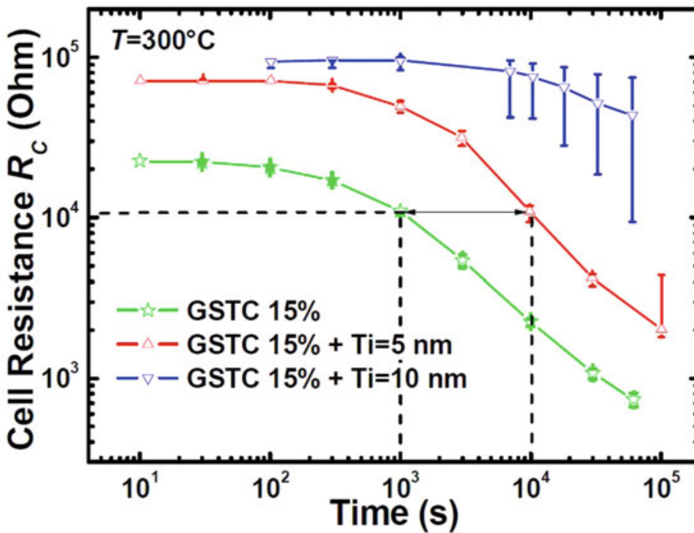


Fig. 7.16 Variation with time of the resistance of virgin GSTC 15% devices (T_{max} 200 °C), for different thickness of the titanium capping layer (© 2016 IEEE. Reprinted, with permission, from [42])

reports, for example, the result of the thermal stability evaluated at 300 °C for the high-resistance state obtained after fabrication (T_{max} 200 °C), confirming the better thermal stability of the devices including a thin-layer titanium.

Despite the improvement in the thermal stability resulting from the addition of carbon and a thin titanium layer, the RESET state of this type of device is not still

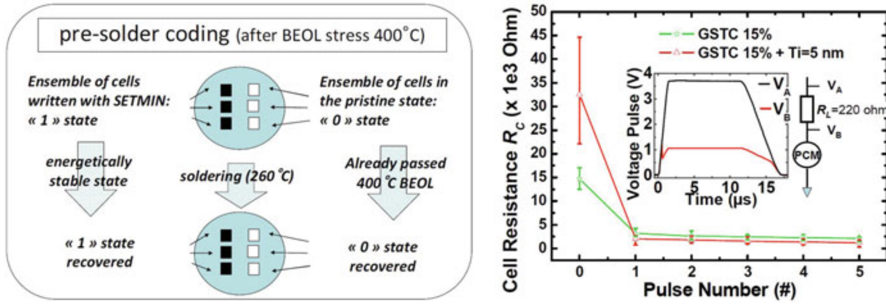


Fig. 7.17 Schematic diagram of the pre-encoding method based on the high-resistive A-AD state obtained at the end of the fabrication, and the low-resistance SET-MIN state obtained when programming the device with a series of pulse as described on the right-hand side graph (© 2016 IEEE. Reprinted, with permission, from [42])

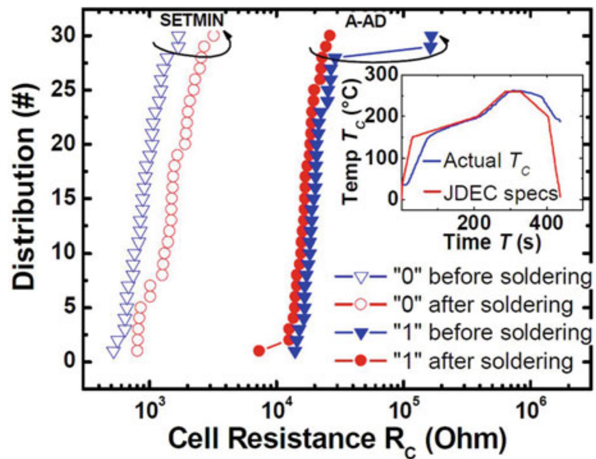


Fig. 7.18 Resistance distribution for the high-resistance state "A-AD" and the low-resistance state "SET-MIN," before and after the application of a thermal profile representative of the soldering step. (© 2016 IEEE. Reprinted, with permission, from [42])

sufficiently stable under the thermal budget applied during the soldering step. An alternative solution to enable the precoding of the information was therefore proposed (Fig. 7.17). If the RESET state is not stable enough to sustain the soldering profile, the state obtained at the output of the fabrication process that is called A-AD (for "amorphous as-deposited state") has a resistance which is significantly higher than that of the crystalline state. Besides, it has the further advantage of being stable with respect to soldering conditions during 2 min at 260 °C since having already undergone the annealing during 1 h at 400 °C during the fabrication. This state can be used as the pre-encoded state of high resistance. A low-resistance state, which is called the SET-MIN state, is programmed by applying a series of electrical pulses as shown in Fig. 7.17**.** Figure 7.18 shows the resistance distributions for the A-AD and SET-MIN states before and after undergoing a thermal stress equivalent to that of the soldering conditions: the resistance of the A-AD state

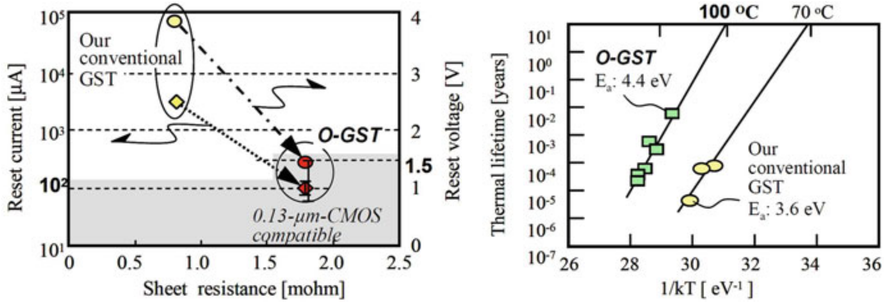


Fig. 7.19 Reset current and voltage versus the sheet resistance of GST and GSTO materials, and data retention of cells in the RESET state at various temperature. (© 2016 IEEE. Reprinted, with permission, from [44])

is stable, while the resistance of the state SET-MIN increases only slightly with a drift coefficient $\nu = 0.02$, leaving open a reading window of more than one order of magnitude between these two states.

Thus, taking advantage of the increased thermal stability of the GSTC alloys, an innovative solution to pre-encode the information before soldering the PCM chip on the electronic board has been proposed. However, we must emphasize the major limitation of this method which is the inability to test the functionality of the memory cells (EWS test for "Electrical Wafer Sorting") before precoding the information. The method relies in fact on the use of state A-AD which is virgin of electrical stress.

7.3.2.3 Oxygen

Oxygen doping was successfully experimented in $Ge_2Sb_2Te_5$, allowing to decrease the RESET current and still maintaining a 10-year data lifetime at 100 °C [44]. Programming currents as low as 100 μA were obtained with a bottom electrode contact of diameter 180 nm (Fig. 7.19). Like in the case of carbon addition, the reduction of the programing current is attributed to the increase of both the electrical and thermal resistances of the material.

7.3.3 Composites Materials Based on the Reference Compounds

Investigations of nanocomposite PCM alloys have successfully been performed, based on the addition of SiO_2 inclusions into GeSbTe. The nanocomposite material was obtained by co-sputtering. On simple pore devices, Czubatyj et al. first demonstrate that SiO2 inclusions exhibit the advantage of lowering the programming

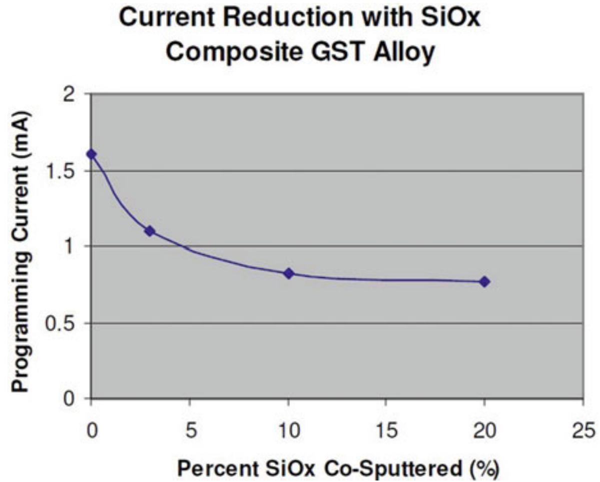


Fig. 7.20 Programming current reduction for varying co-sputtered percentages of SiO_2 with GST material. Reprinted with permission from [45]

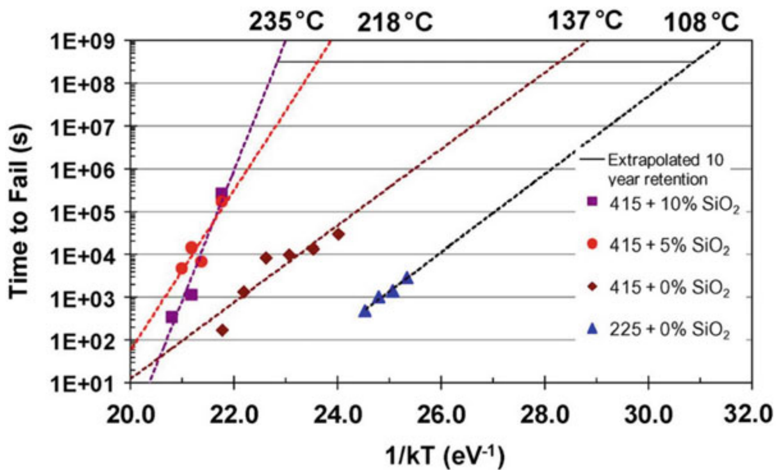


Fig. 7.21 Data retention of the RESET state highlighting the higher retention forecasted at 10 years for GST + SiO_2 devices (© 2016 IEEE. Reprinted, with permission, from [46])

current, resulting from the increase of both the electrical and thermal resistivity of the devices [45]. From thin-film resistivity measurements, the crystallization temperature is shown to increase by about 70 °C for 15% SiO_2 added into $Ge_4Sb_1Te_5$. The better thermal stability of the amorphous phase results in an increase of the data (Fig. 7.20) retention for the RESET state of the corresponding devices (Fig. 7.21). The data retention is in fact extrapolated to 10 years at 235 °C for 10% SiO_2 [46], so that the composite material fits both the specification for the automotive applications and for the pre-encoding of the data at the wafer level before soldering. The performances of the nanocomposite PCM have however not yet been demonstrated with statistical data.

7.4 Optimization of PCM Devices Integrating Nonstoichiometric Alloys from the GeSbTe System

To further enhance the PCM device performances, GST alloys with different compositions from the reference stoichiometric compounds have been investigated. As an example, Lee et al. present the results obtained with 1Gb cell arrays showing an improved thermal stability with respect to $Ge_2Sb_2Te_5$ [2]: the data retention is extrapolated to 10 years at 200 °C and a SET programming time of 300 ns is demonstrated. While the composition of the alloy $Ge_xSb_yTe_z$ is not disclosed, more precise investigations performed throughout the GST system reveal that Ge-rich GST alloys can target high-temperature applications, while Sb-rich GST alloys are well suited for high-speed performances (Fig. 7.22).

7.4.1 Ge-Rich GST Alloys

In 2011, Cheng et al. paved the way for the study of the GeSbTe ternary system toward Ge-rich compositions [47], which they hereafter further optimized with the addition of nitrogen [48]. Focusing their attention to the pseudo-binary line Ge-Sb_2Te_3, the high performance of the so-called "golden" composition was highlighted. This germanium-rich composition was shown to particularly enhance the thermal stability of the RESET state while maintaining a fast switching speed. Using Ge-rich compositions, Zuliani et al. were able to demonstrate the data retention during the typical soldering reflow temperature profile [49]. Meanwhile,

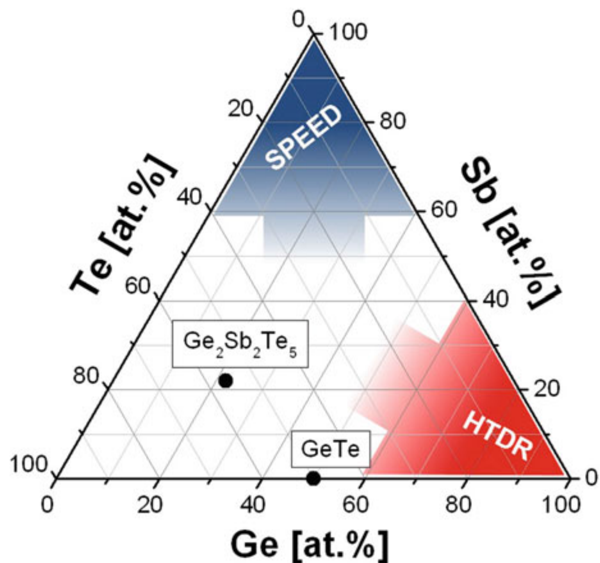


Fig. 7.22 Schematic representation of the GeSbTe material engineering for PCM device optimization

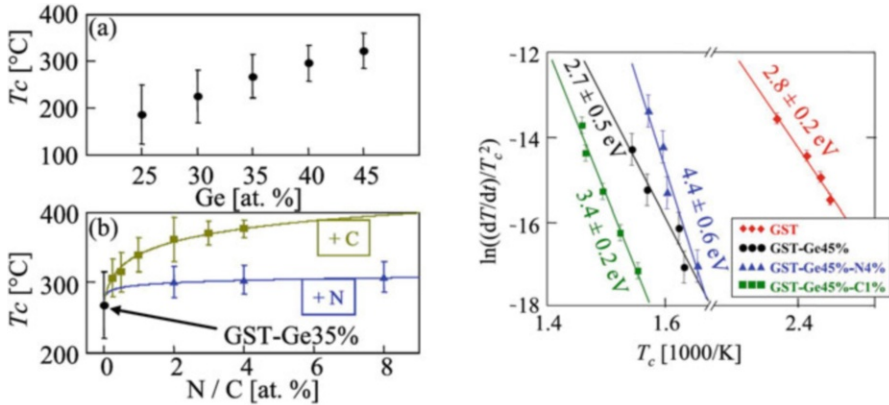


Fig. 7.23 Crystallization temperature and activation energy for crystallization of Ge-rich GST alloys, as a function of the germanium percentage added into $Ge_2Sb_2Te_5$. Nitrogen and carbon additions are also considered (© 2016 IEEE. Reprinted, with permission, from [18])

through the study of the pseudo-binary line Ge-$Ge_2Sb_2Te_5$, an interpretation was proposed for the high stability of both the RESET and SET states [18, 50]. In the following, we report the results of these investigations. The notation "GST + Ge x %" refers to the nominal fraction of Germanium added to the $Ge_2Sb_2Te_5$ alloy using co-sputtering from two targets, namely, Ge and $Ge_2Sb_2Te_5$.

At the film level, X-ray diffraction patterns measured on Ge-rich GST thin films after annealing at 400 °C show the presence of two cubic phases corresponding to the separation of Ge and $Ge_2Sb_2Te_5$ phase indicated in the equilibrium phase diagram [51]. The volume fraction of the germanium phase is, as expected, all the more important as the germanium percentage is high. For this type of germanium-rich GST alloys and unlike the stoichiometric compounds $Ge_2Sb_2Te_5$ or GeTe presented before, the crystallization involves a local composition change accompanying the phase separation. One can note that for equal germanium content, the volume fraction of the germanium cubic phase is reduced due to the addition of nitrogen, similar to the results observed for GeTeN alloy, suggesting that a part of germanium is trapped by the nitrogen in the formation of Ge-N bonds.

The results of the thin-film characterizations confirm the greater thermal stability of the amorphous phase with the germanium enrichment, which is illustrated by the increase in crystallization temperature (Fig. 7.23). The addition of light elements such as nitrogen or carbon reinforces this stability, as shown by both the increase in temperature of crystallization and also that of the activation energy for crystallization.

Germanium-rich GST alloys have been incorporated into an industrial test vehicle, comprising PCM memory cell structure "wall" [49] and manufactured on a 90 nm CMOS platform. Despite the phase separation observed when characterizing the thin films, the devices exhibit functional RI characteristics (Fig. 7.24): these RI characteristics show that the SET to RESET transitions are not as abrupt as

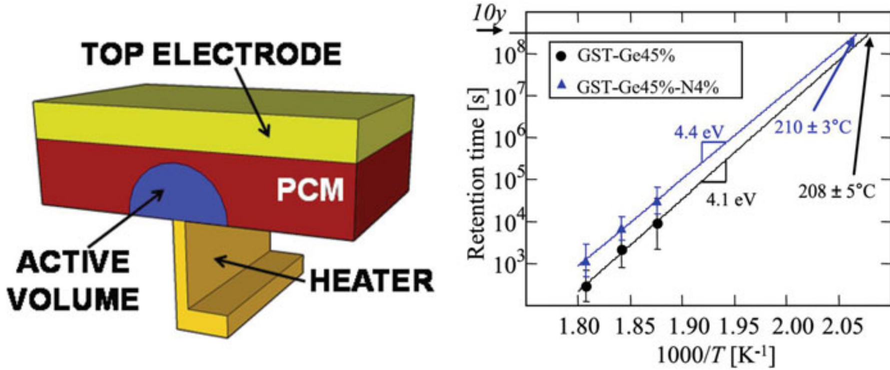


Fig. 7.24 Schematic drawing of the PCM wall storage element and data retention of the RESET state showing the extrapolation to 10 years (© 2016 IEEE. Reprinted, with permission, from [18])

the one obtained with the stoichiometric compounds $Ge_2Sb_2Te_5$, but the RESET current is substantially unchanged. With respect to $Ge_2Sb_2Te_5$ devices, the highest thermal stability of the amorphous phase is illustrated by the better retention of the devices programmed in the RESET state: for the GST alloys enriched with 45% germanium, the extrapolation of the measurements indicates the stability of the RESET state for at least 10 years at temperatures up to 200 °C, well above the required specification for automotive applications (10 years at 150 °C).

Having demonstrated the excellent thermal stability of the RESET state for the devices integrating GST enriched with germanium, the stability of the SET state is next evaluated. The SET state shows a resistance increase over time induced by the structural relaxation of residual disordered regions such that grain boundaries [52]. This resistance drift has to be minimized because it may lead to the loss of the low resistance (SET state). If the resistance drift of the SET state of the reference phase change materials such as $Ge_2Sb_2Te_5$ is negligible, Fig. 7.25 shows that it is highly accentuated by the addition of germanium. Figure 7.26 shows on the other hand the link between the value of the SET resistance and the value of its drift: the lower the SET resistance, the lower the SET drift.

To minimize the drift of the SET state, a programming procedure able to minimize the SET resistance level and hence the drift of the SET state was proposed. Figure 7.27 shows that the programming procedure "R-SET", consisting of a RESET pulse immediately followed by a SET pulse of low intensity, allows to achieve lower SET resistance than those obtained by the standard single-pulse SET procedure performed at higher intensity. This effect results from the decrease of the threshold voltage V_{th} of the phase change material immediately after the application of the RESET pulse [53]. The lower threshold voltage V_{th} allows to program the devices in the SET state using lower voltages and thus lower currents that prove to be more favorable to crystallization.

Figure 7.28 confirms, as expected, that the drift of the devices programmed with the "R-SET" procedure is lower than that of devices obtained with the standard SET

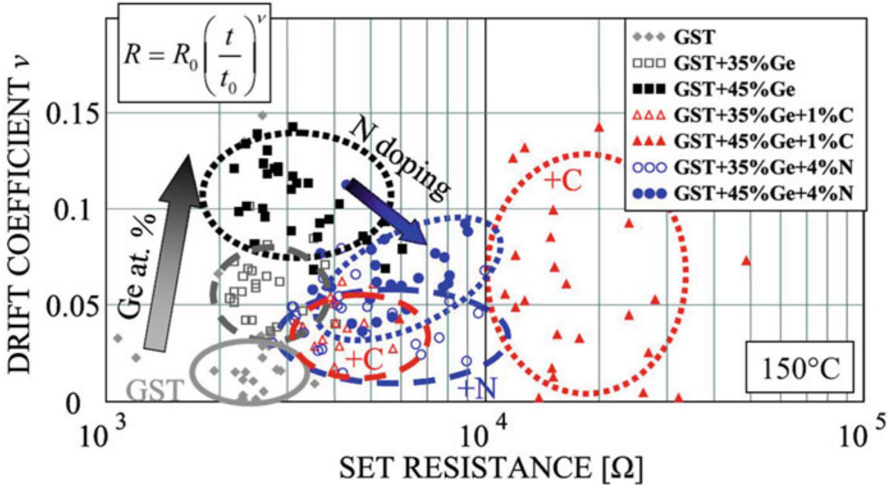


Fig. 7.25 Drift coefficient for the low-resistance state, evaluated with different Ge-rich GST materials. The cells have been programmed to the minimum resistance state by the application of a series of pulse of decreasing amplitude ("SET-SCD" procedure) (© 2016 IEEE. Reprinted, with permission, from [18])

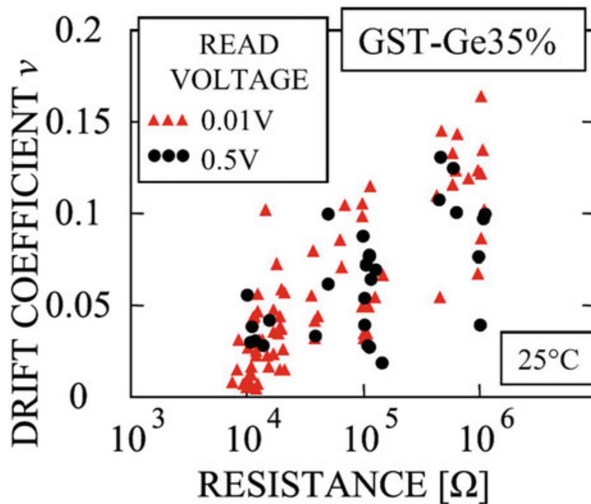


Fig. 7.26 Drift coefficient as a function of the programmed resistance of the GST + Ge 35%-based devices (© 2016 IEEE. Reprinted, with permission, from [18])

procedure. The figure also reports for comparison the results obtained with the application of a series of pulses of decreasing amplitude which is referred to as the SET-MIN procedure. This procedure allows to obtain the SET-MIN state of minimum resistance and drift, but the programming time (~15 μs) is too long to guarantee competitive program bandwidth with respect to other nonvolatile memory technologies.

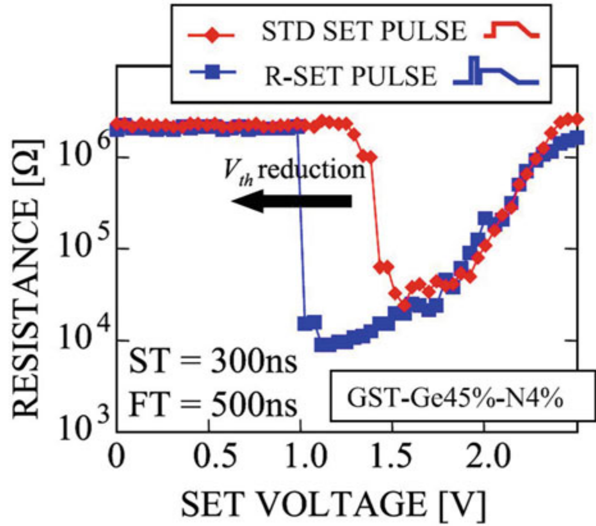


Fig. 7.27 Programming characteristics of GST + Ge devices as obtained by single trapezoidal programming pulses (“STD-SET”), and eventually preceded by a RESET pulse (“R-SET”) (© 2016 IEEE. Reprinted, with permission, from [18])

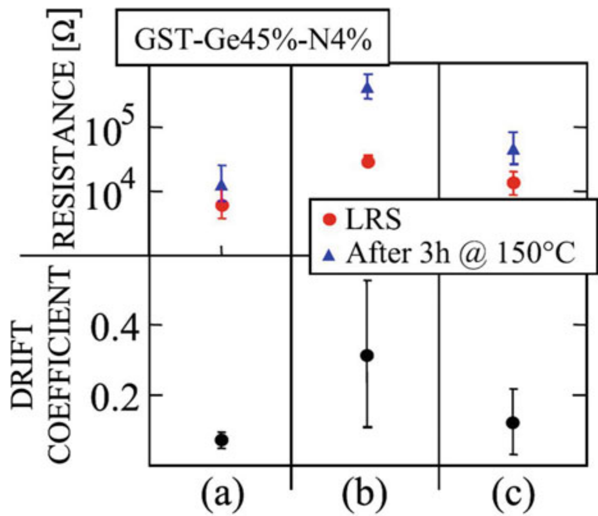


Fig. 7.28 Programmed resistance and drift coefficient for various low-resistance states, programmed (**a**) with a series of pulses of decreasing amplitude (“SET-SCD”), (**b**) with a single trapezoidal pulse of high intensity (“STD-SET”), (**c**) with a RESET pulse immediately followed by a trapezoidal pulse of low intensity (“R-SET”) (© 2016 IEEE. Reprinted, with permission, from [18])

As already pointed out, the hardest solicitation during which a PCM memory must retain programmed information is the soldering step of the component on the printed circuit board. The stability of the SET and RESET states of germanium-rich GST devices was verified when the devices are subjected to a temperature profile similar to the annealing performed during the soldering step (RSTP annealing for “Reflow Soldering Temperature Profile”) and including a peak temperature at 260 °C for 2 min. Figure 7.29 shows that the two programmed states maintain a resistance window of more than one order of magnitude after the RSTP annealing,

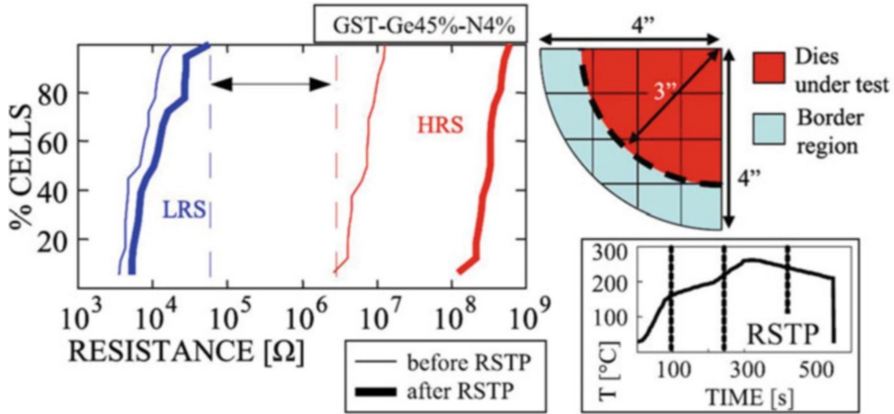


Fig. 7.29 Resistance distribution before and after the application of a reflow soldering thermal profile (RSTP) (© 2016 IEEE. Reprinted, with permission, from [18])

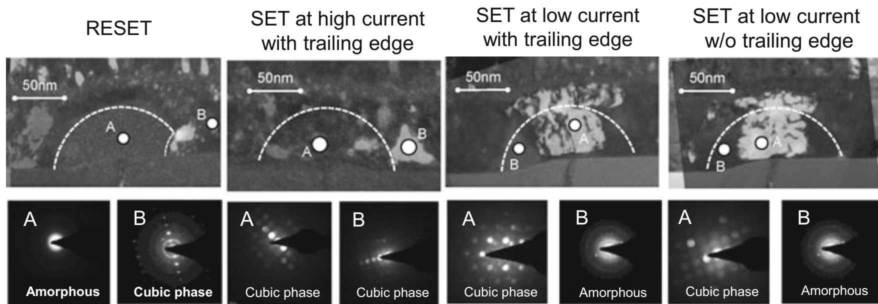


Fig. 7.30 Superimposition of bright and dark field TEM images for GST + Ge 45% devices programmed to different states

which validates this type of materials for embedded applications requiring pre-programming of the memory components at the wafer level.

The functionality of the germanium-rich GST devices and even more the endurance of this type of devices shown to reach more than 10^7 cycles [49] may be surprising at first sight, given the observed phase separation at the film level. To consolidate the electrical results, a thorough morphological analysis was conducted by transmission electron microscopy on devices programmed into different states (Fig. 7.30). The morphological analysis was supplemented by numerical simulations, to interpret the experimental images (Fig. 7.31).

- RESET state: as expected, an amorphous dome overlooking the "heater" element is identified on the image.
- SET state obtained with a trapezoidal pulse of high intensity: the phase change material layer appears to be polycrystalline and several grains can be imaged in

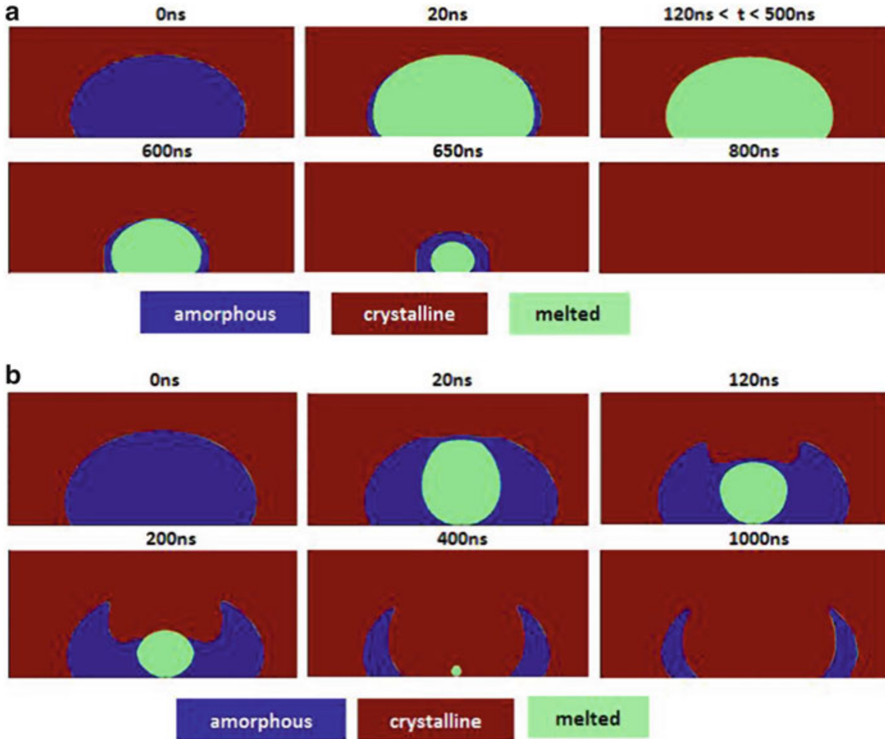


Fig. 7.31 Simulation of the phase change during SET programming, performed at (**a**) high intensity and (**b**) low intensity. The pulse comprises a plateau of duration 500 ns followed by a trailing edge (© 2016 IEEE. Reprinted, with permission, from [54])

the thickness of the layer. The simulation shows that this polycrystalline structure results from the following sequence: first, during the plateau of the trapezoidal SET pulse, complete melting of the amorphous dome occurs; and second, during the descent of the pulse, the recrystallization of the melted area proceeds by the growth of the crystal grains located at the periphery of the amorphous dome.

- SET state obtained by a trapezoidal pulse of low intensity (“R-SET” procedure): the TEM images and the associated diffraction patterns show the presence of a crystalline column passing through the amorphous dome. The dark field images reveal the unique crystal orientation of this column, thus highlighting the absence of grain boundaries along the conductive path in the R-SET state. The simulation of this programming pulse shows that this peculiar structure results from the following mechanism. First, the amorphous phase switches to the conductive ON state along a filament which passes through the dome amorphous, thus leading to the partial melting of the material phase change along this axis. Then, the temperature rise at the top of amorphous dome leads to the growth of the crystal grain, which gradually proceeds toward the inside of the

dome in the direction of the heater element. This dynamic phenomenon continues for about 300 ns during the plateau of R-SET pulse, until a steady state is reached and the growth of the column is stopped. The situation is such that at the end plate of the R-SET pulse, the amorphous dome is crossed through by a crystalline column, whereas a residual molten zone of reduced size (few nanometers) persists above the heater element. It is interesting to note that the residual amorphous area can be imaged by TEM when observing a device programmed with a square-shaped R-SET pulse at low current. In contrast, no amorphous residue is visible when the R-SET pulse includes a decay time: in this case, the residual liquid crystallizes during the falling edge of the pulse, as confirmed by simulations.

- SET-MIN state obtained by a series of pulses of decreasing amplitude: the phase change material appears to be polycrystalline with some large grains located in the active area of the device. This result is interpreted as the grain ripening induced by the successive overheating suffered at the active area during the programming procedure.

The comparison of the morphology of the different programmed devices allows to correlate therefore qualitatively the drift of the resistance to the presence of grain boundaries along the conductive path: the higher the number of grain boundaries along the conductive path, the higher is the resistance drift of the device. This observation reinforces the interpretation proposed by Ciocchini et al. [52] according to which the drift observed with devices programmed in the SET state results from the structural relaxation of the residual amorphous regions, including those located at the grain boundaries of the polycrystalline material. From an application point of view, this study also highlights that the resistance drift of the SET state can be minimized by programming the device with a low current intensity.

The analysis of the elemental distribution performed by EELS mapping on such devices shows the strong inhomogeneity in composition of the active region for the devices programmed in the SET and RESET states (Fig. 7.32): the core of the active zone located just above the heater element is highly enriched with antimony, whereas the side areas appear highly enriched with germanium. The overlay of the images showing the morphology of devices and those showing the local composition reveals that the germanium-rich areas are crystalline. The inhomogeneity of the elemental composition appears overall stable throughout the memory cell programming cycles. It is attributed to a germanium segregation phenomenon [55]. The fact that the germanium depletion extends to the upper electrode, that is to say beyond the active volume which is restricted to the amorphous dome during the write/erase cycles, suggests that the segregation phenomenon largely occurs during the first activation of the devices. Indeed, after the manufacturing process, the germanium-rich GST devices are in a high-resistance state, because they consist of a mixture of fine GST and Ge grains. The application of a voltage equivalent to that of a RESET pulse results in the formation of a conductive path between the heater element and the upper electrode, along which the phase change material gets

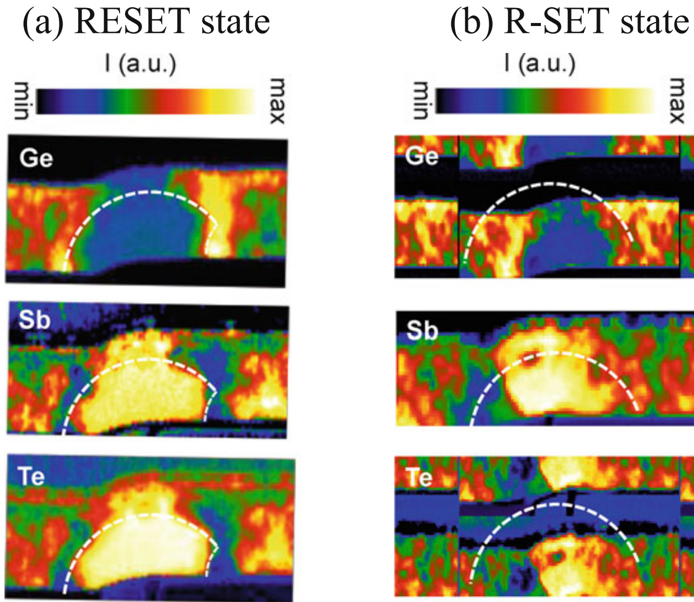


Fig. 7.32 STEM/EELS images of GST + Ge 45% devices showing the Ge, Sb, and Te distributions in the active part of the PCM storage element (© 2016 IEEE. Reprinted, with permission, from [50])

melted. The periphery of the molten zone can then be seen as a germanium well due to the segregation effect: germanium solidifies in the temperature range 640–920 °C, the solid germanium phase being in equilibrium with a liquid phase, which is a GeSbTe alloy comprising less germanium than the initial composition of the device. The high germanium content in the nominal alloy GST + Ge 45at% allows to save, after this segregation effect inherent to the cell activation, an alloy equivalent to GST + Ge 25% at the core of the cell, thus ensuring the thermal stability of the RESET state at temperatures up to about 250 °C. Note that this maximum temperature able to ensure the RESET state stability is very close to the crystallization temperature of the alloy GST + Ge 25% identified at the core of the devices (Fig. 7.23). The consistency of these results consolidates the proposed interpretation.

The performance of these new germanium-rich GST materials was finally validated by the integration in a 12 Mb test vehicle comprising state-of-the-art PCM storage elements [49]. The RESET programming current of these small devices is of the order of 400 μA. The programming cartographies confirm that a low intensity pulse (plateau 300 ns, intensity 220 μA, and trailing edge 400 ns) minimizes the value of the SET resistance and consequently the drift coefficient (Fig. 7.33). Under the application of a series of successive anneals for a period of 1 h at increasing temperatures, the decrease of the resistance of the RESET state resulting from the effect of the crystallization is observed from 260 °C, whereas the

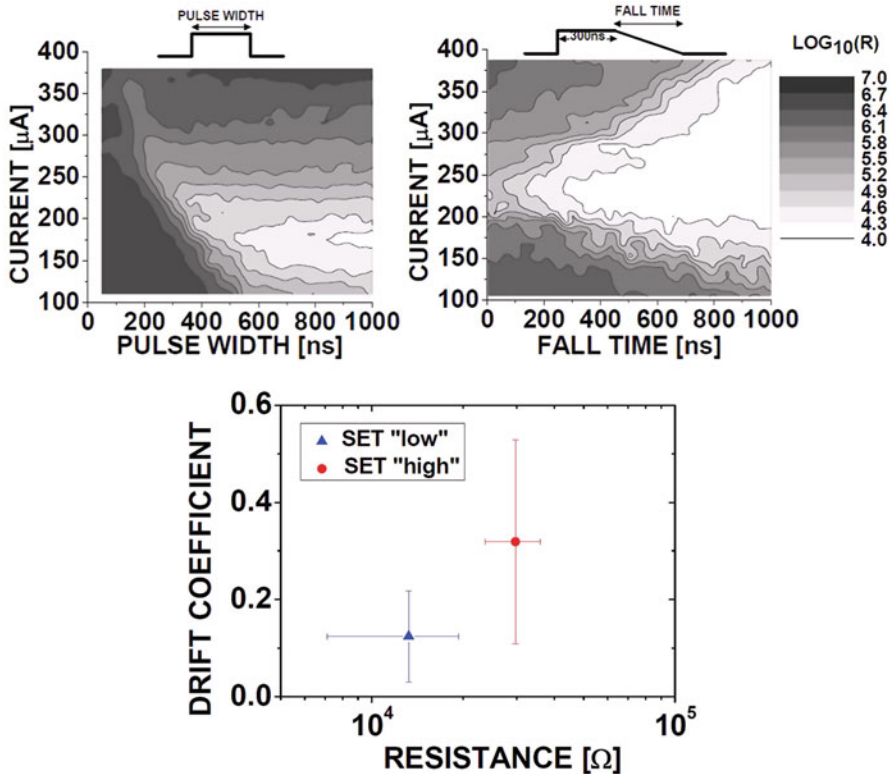


Fig. 7.33 Programming cartographies showing that a pulse of low intensity ~220 µA with a plateau of 300 ns and trailing edge of 300 ns allows to minimize the resistance of the device and hereafter the drift coefficient (© 2016 IEEE. Reprinted, with permission, from [50])

SET state of low resistance is stable. The distributions of the SET and RESET resistances assessed on the 12 Mb memory array show that the reading window is preserved after annealing for 2 h at 230 °C and in the whole range of studied content, or ±2at% about the composition of optimized alloy, thus illustrating the wide process window associated with this alloy (Fig. 7.34). The reading window is preserved after at least 107 write/erase cycles, and the crystallization temperature of the active material is seen reduced by only 10 °C after 10^7 cycles (Fig. 7.35).

Thus, it has been demonstrated that Ge-rich GST alloys meet the specifications required for the embedded memory market, allowing to achieve the compromise targeted between the programming speed, the retention of the RESET state, and the drift of the SET state. More precisely, Zuliani et al. [56] point out that the Ge content in Ge-rich GST alloys can be tuned according to the targeted applications. Two different compositions are highlighted. The two alloys are both compatible with automotive applications. On top of this, the first alloy fits with the most challenging requirement of data retention after soldering reflow, while the second

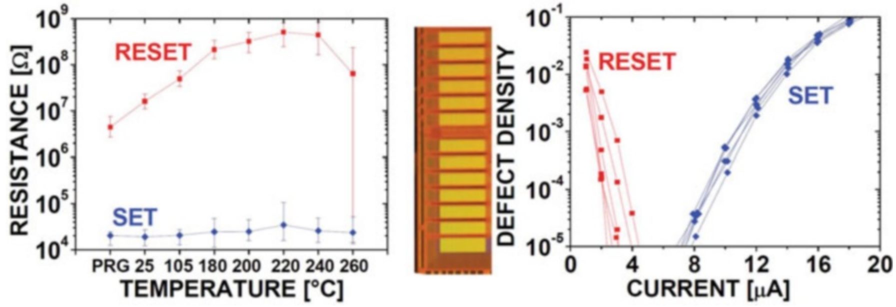


Fig. 7.34 Performance of a Ge-rich GST device, illustrating the high thermal stability of the programmed states. (**a**) SET/RESET resistances evaluated on analytical cells after successive 1 h anneals up to 240 °C. (**b**) SET/RESET distributions evaluated on the 12 Mb demonstrator after a 2 h anneal at 230 °C and for various germanium contents (±2at% from the optimized alloy) (© 2016 IEEE. Reprinted, with permission, from [50])

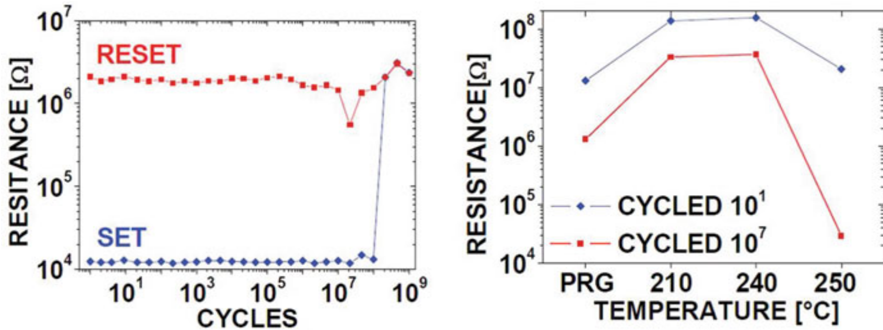


Fig. 7.35 Performance of Ge-rich GST devices illustrating the endurance of the devices (© 2016 IEEE. Reprinted, with permission, from [50])

one exhibit better properties on the SET side (faster crystallization, lower SET resistance, and higher reading window).

7.4.2 Sb-Rich GST Alloys

As opposite to the Ge-rich GST compositions which are investigated for the high thermal stability of the amorphous material with respect to crystallization, the Sb-rich corner of the GST system primarily attracts attention for the high crystallization speed it offers. In fact and as known from material investigation originally dedicated to optical discs, Sb-rich materials such as doped $Sb_{69}Te_{31}$ [57] or $Ge_{15}Sb_{85}$ [58] exhibit a fast crystallization, which involves a growth-dominated mechanism. Investigations of the $Ge_1Sb_xTe_1$ alloys ($x > 1$) further show that the crystallization time of the alloy $Ge_1Sb_6Te_1$ can be as low as 20 ns. Meanwhile, the

crystallization temperature is substantially higher than that of the reference compound $Ge_2Sb_2Te_5$. Thus, Sb-rich alloys appear as promising materials for storage class memories which require high-speed operation. Kim and coworkers [59] even demonstrated a confined cell 7.5 × 17 nm whose high speed of operation (20 ns), high endurance (extrapolated to 10^{15} cycles), and data retention (4.5 years at 85 °C) allowed them to envisage the PCM technology for DRAM-type applications (Fig. 7.36).

7.4.3 Te-Poor GST Alloys

To complete the investigation of the GST system, Boniardi and coworkers present the performance obtained with the Te-poor family (Fig. 7.37), targeting intermediate compositions between the Ge-rich and the Sb-rich alloys previously presented [61, 62]. Their motivation to go in the direction of Te-poor compositions was supported by the increase of the average coordination number of the alloys, likely having a major impact on the crystallization properties. In fact, a decreasing tellurium content is known to result in an increased rigidity of the alloy [63], thus likely to favor a higher stability of the amorphous phase over time. The RESET resistance decreases with decreasing Te content, thus resulting in a smaller reading margin. The decrease of the RESET resistance is attributed to the decrease of the mobility gap, as illustrated by the lower activation energy for conduction. The decrease of the mobility gap also results in a decrease of both the threshold switching and holding voltage. As expected, Te-poor compositions result in higher SET and RESET currents. The analysis of the injected power required to reach the melting condition as a function of the external temperature allows to evaluate the thermal resistance of the cell. The decrease in Te content is accompanied by a decrease in the thermal resistance of the cell, thus highlighting the main contribution at the origin of the higher programming current (Figs. 7.38 and 7.39).

7.5 Optimization of PCM Devices Based on Other Material Systems

Besides the GeSbTe ternary system, research efforts have also been dedicated to the development of phase change materials outside of this reference system. At the material level, the investigations first target the identification of compositions with a high electrical contrast likely to result in a large reading window and with a high crystallization temperature to ensure a high thermal stability of the amorphous state. Few studies further try to validate the proposed material with the characterization of the device performances, targeting a high programming speed, a low programing current, and a high endurance. Recently, however, remarkable results

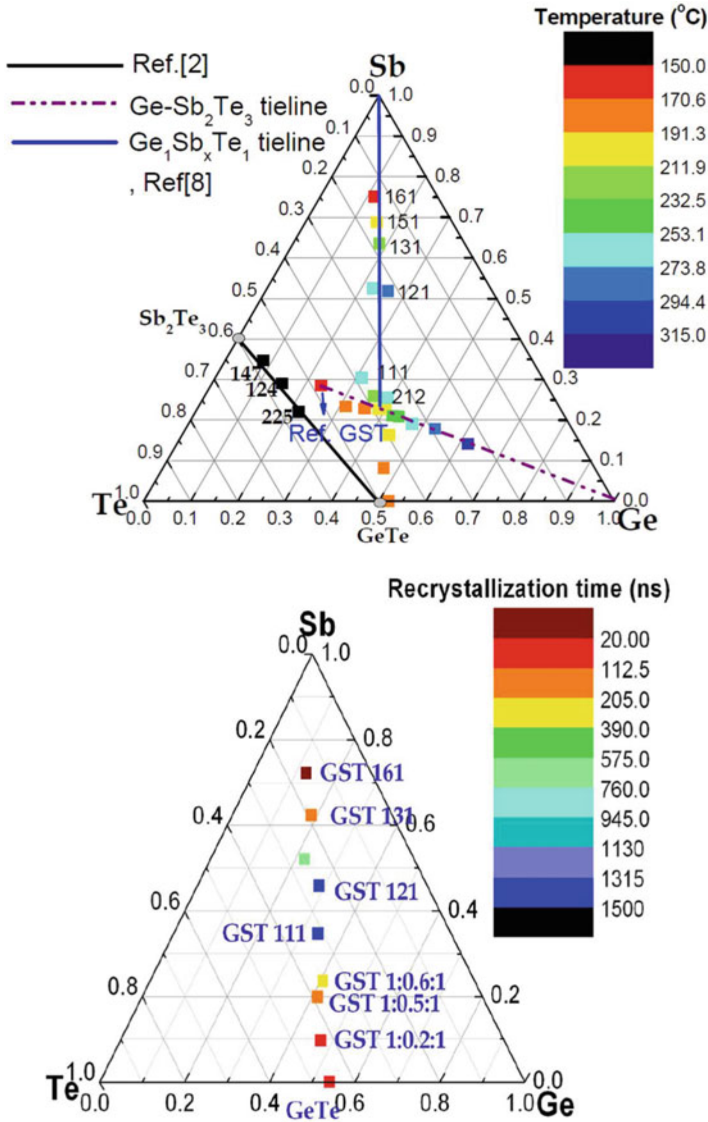


Fig. 7.36 Crystallization temperature (© 2016 IEEE. Reprinted, with permission, from [47]) and recrystallization times (Reprinted from [60], Copyright 2016, with permission from Elsevier) of $Ge_1Sb_xTe_1$ materials showing the necessary trade-off between the stability of the amorphous phase and the high crystallization speed of Sb-rich GST alloys

have been presented by Cheng and coworkers, who proposed to use doped GaSbGe [64]. The nature of the dopant element is not revealed, but the alloy is said to lie on the pseudo-binary tie-line Ge – $Ga_{46}Sb_{54}$. Figure 7.40 highlights the increasing crystallization temperature with increasing Ge content, thus approaching 400 °C for

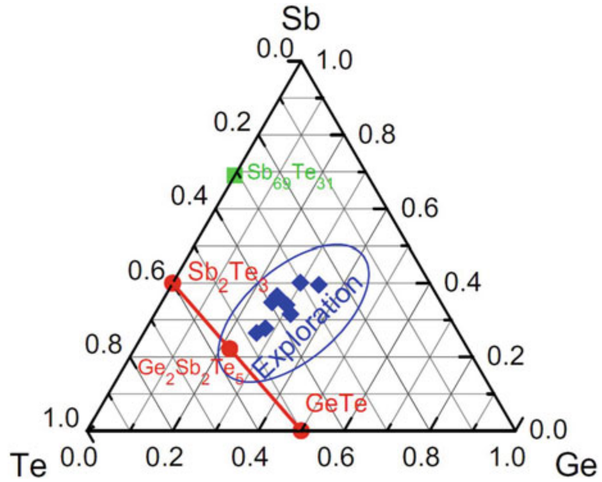


Fig. 7.37 The GeSbTe ternary compound system and the exploration in the Te-poor region performed by Boniardi and coworkers. Reprinted from [61], Copyright 2016. with permission from Elsevier

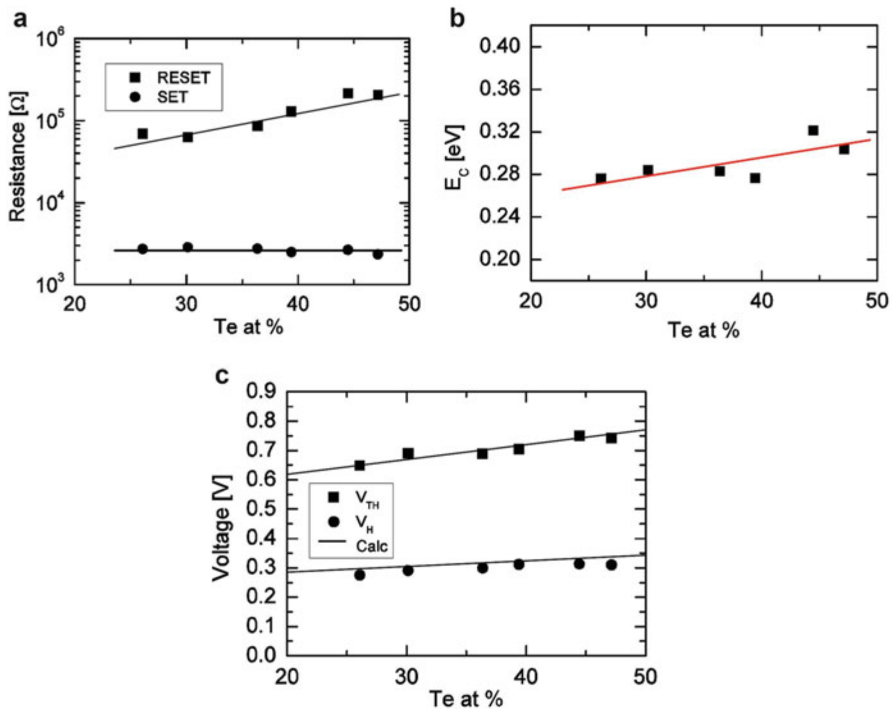


Fig. 7.38 Variation of the electrical characteristics of Te-poor GST devices studied by Boniardi and coworkers as a function of the Te content (**a**) set and RESET resistance, (**b**) activation energy of conduction E_c for the RESET state, (**c**) threshold switching voltage V_{th} and holding voltage V_h for the RESET state (© 2016 IEEE. Reprinted, with permission, from [62])

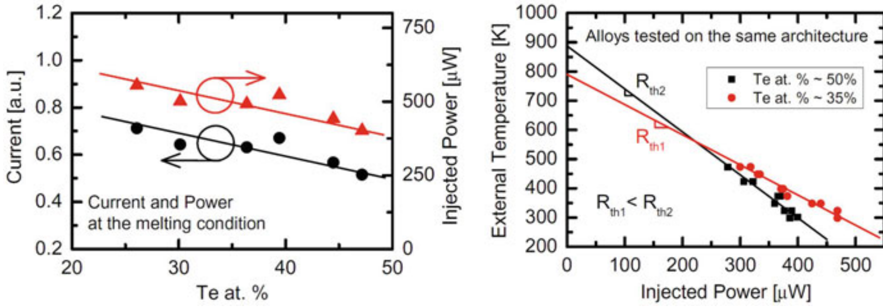


Fig. 7.39 Programming current and mower to reach the melting condition and thermal resistance extraction as a function the Te at% (Reprinted from [61]. Copyright 2016, with permission from Elsevier)

the selected alloy. This alloy was further integrated in a 128 Mb PCM test vehicle with TiN ring bottom electrodes of 20–40 nm diameter. Note that the wafers were submitted to a 525 °C post-fabrication anneal for 2 min to enable the crystallization of the phase change material. The electrical results displayed in Fig. 7.41 show that the alloy enables to successfully pass the solder reflow test considering the two programmed states. Besides, the data retention can be extrapolated to 10 years at 220 °C. Even though elemental segregation is observed in the programmed devices, an endurance of 10^9 cycles is successfully demonstrated. As far as the programming speed is concerned, a one-decade reading window is achieved using SET programming pulses as short as 80 ns (Fig. 7.42). Moreover, the low values of the SET and RESET resistance states, which are about one order of magnitude lower than that obtained with GST-based materials, can enable a fast reading operation. Thus, GaSbGe is the fastest material that can pass the soldering bonding criteria to date. The main drawback of this new material appears to lie in the relatively high SET programming current, and no results on the RESET current have yet been published. This investigation shows that there is probably some room for new material systems besides GST to be investigated for PCM devices.

7.6 Conclusions

Phase change material engineering can be cleverly utilized to optimize the performances of PCM devices. However, because of some conflicting parameters like the thermal stability of the programmed states and the programming speed, it appears that the phase change materials need to be properly customized for a given application.

Stoichiometric compounds like $Ge_2Sb_2Te_5$ and GeTe allow short programming times (in the order of a few tens of nanoseconds), but the thermal stability of the programmed states is limited (about 10 years at 100 °C) due to the low

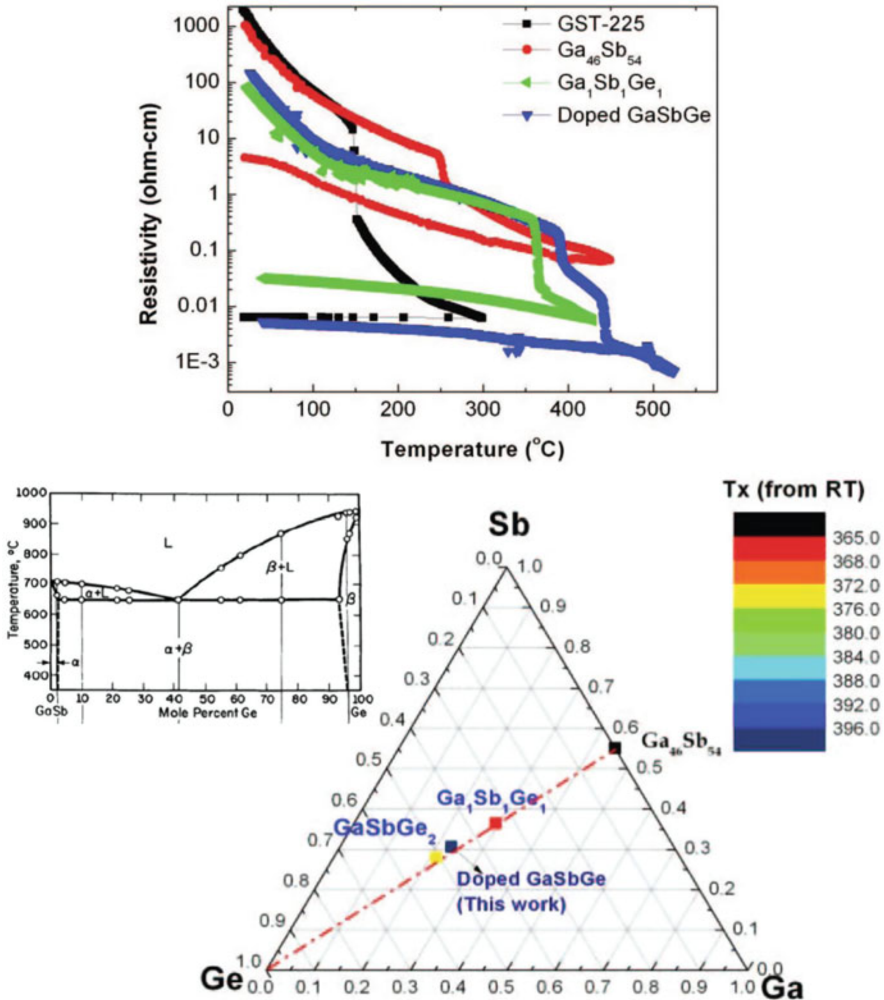


Fig. 7.40 Resistivity as a function of the temperature and crystallization temperature Tx as a function of compositions in the Ga-Sb-Ge ternary phase diagram. Tx is increased by Ge and dopants incorporation along the Ge/GaSb tie-line. The inset shows the GaSb-Ge pseudo-binary phase diagram from [53]. © 2016 IEEE. Reprinted, with permission, from [64]

crystallization temperature of the amorphous phase. The addition of light elements into these stoichiometric alloys allows extending the temperature range in which the programmed states are stable: for example, the addition of nitrogen into GeTe or carbon into $Ge_2Sb_2Te_5$ allows to achieve the specification required for automotive applications (at least 10 years at 150 °C). However, the hardest specification that is keeping the programmed states during the soldering step (2 min at 260 °C) has not been shown to date with phase change materials based on the stoichiometric compounds.

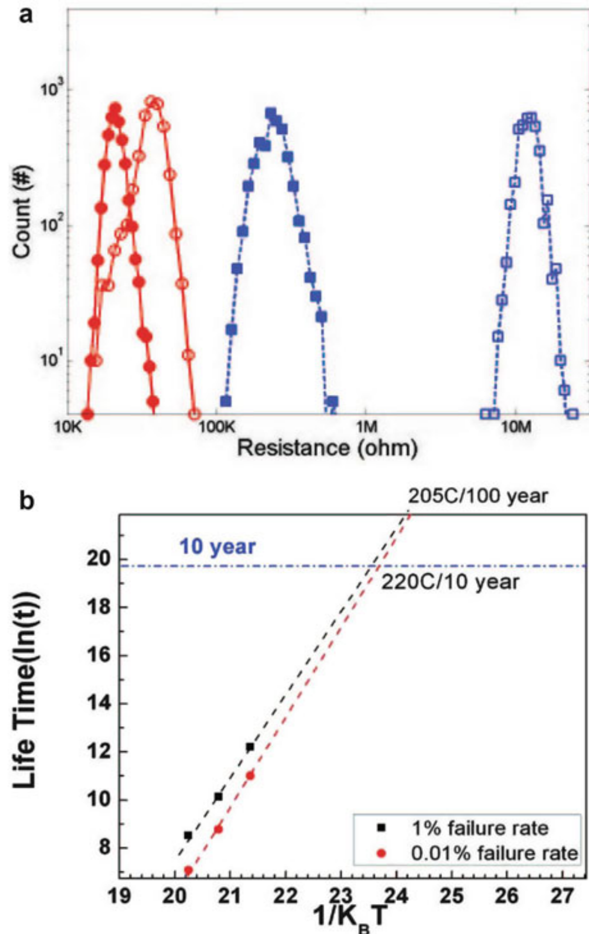


Fig. 7.41 (**a**) SET and RESET resistance distributions before (*closed symbols*) and after (*open symbols*) solder bonding reflow thermal budget for 4 k cells in a 128 Mb test chip. (**b**) Data lifetime for doped GaSbTe showing the extrapolation of the data retention to 10 years at 220 °C (© 2016 IEEE. Reprinted, with permission, from [64])

The investigation of nonstoichiometric GST alloys allows to further extend the performance of the PCM devices: the addition of germanium tends to increase the thermal stability of the programmed states, while adding antimony ensures a very high programming speed (Fig. 7.22). Thus, the stability of the programmed states under the soldering reflow conditions which is necessary for some embedded applications was demonstrated with Ge-rich GST alloys. On the other hand, Sb-rich GST alloys are forecasted for being used in storage class memories.

While the GST system has been thoroughly investigated, other material systems today appear promising to further extend the PCM device performances. This opens new ways for future material engineering investigations targeting PCM device optimization.

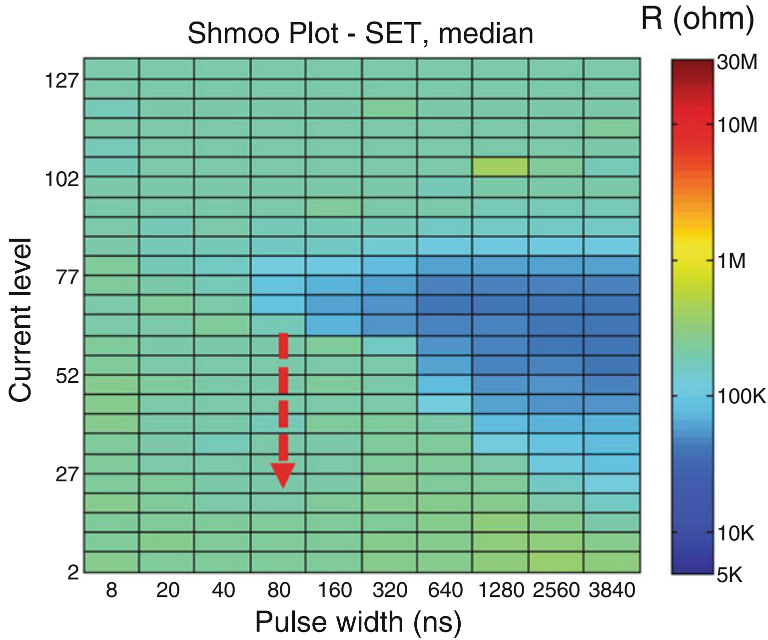


Fig. 7.42 Programmed resistance versus pulse width and intensity for doped GaSbGe PCM cells integrated in a 128 Mb test chip, showing that a 10× resistance ratio can be obtained with an 80 ns SET programming pulse (© 2016 IEEE. Reprinted, with permission, from [64])

References

1. G. Servalli, A 45nm generation phase change memory technology. IEEE International Electron Devices Meeting (IEDM) Location: Baltimore, MD Date: DEC 7–9, 2009; Published: 2010
2. S.H. Lee, et al., Highly productive PCRAM technology platform: Based on 4F2 (84nm pitch) cell scheme for 1 Gb and beyond. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 5–7, 2011; Published: 2012
3. Y. Choi, et al., A 20nm 1.8V 8Gb PRAM with 40MB/s Program Bandwidth. IEEE International Solid-State Circuits Conference (ISSCC) Location: San Francisco, CA Date: FEB 19–23, 2012; Published: 2012
4. M.J. Kang, et al., PRAM cell technology and characterization in 20nm node size. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 5–7, 2011; Published: 2012
5. R. Krishnamurthy, Inside the Samsung 512 Mb Phase Change Memory at http://ww3.chipworks.com/ko/technical-competitive-analysis/resources/blog/a-peek-inside-the-samsung-512-mb-4f2-cross- point-phase-change-memory/ (2011)
6. R. Krishnamurthy, First volume production phase change memory by micron : A peek inside the micron 1 Gb, 45 nm PCM at http://ww3.chipworks.com/ia/technical-competitive-analvsis/resources/blog/first-volume-production-phase-change-memory-by-mi- cron/ (2013)

7. S.R. Ovshinsky, Reversible electrical switching phenomena in disordered structures. Phys. Rev. Lett. **21**, 1450–1453 (1968)
8. N. Yamada et al., High speed overwritable phase change optical disc material. Jpn. J. Appl. Phys. **26**, 61–66 (1987)
9. Downloaded from the site http://www.ovonic.com (1999)
10. K. Shportko et al., Resonant bonding in crystalline phase-change materials. Nat. Mater. **7**, 653–658 (2008)
11. G. Bruns, P. Merkelbach, C. Schlockermann, et al., Nanosecond switching in GeTe phase change memory cells. Appl. Phys. Lett. **95**, 043108 (2009)
12. A. Kolobov et al., Understanding the phase-change mechanism of rewritable optical media. Nat. Mater. **3**, 703–708 (2004)
13. M. Wuttig, N. Yamada, Phase-change materials for rewriteable data storage. Nat. Mater. **6**, 824–832 (2007)
14. D. Lencer et al., A map for phase-change materials. Nat. Mater. **7**, 972–977 (2008)
15. U. Russo, D. Ielmini, A. Redaelli, A.L. Lacaita, Intrinsic data retention in nanoscaled phase-change memories—Part I: monte carlo model for crystallization and percolation. IEEE Transactions on Electron Devices **53**(12), 3032–3039 (2006)
16. S. Lavizzari, D. Ielmini, A. Lacaita, A new transient model for recovery and relaxation oscillations in phase-change memories. IEEE Trans. Electron Devices **56**(5), 1078–1085 (2009)
17. N. Ciocchini, E. Palumbo, M. Borghi, et al., Modeling resistance instabilities of set and reset states in phase change memory with Ge- rich GeSbTe. IEEE Trans. Electron Devices **61**(**6**), 2136–2144 (2014)
18. G. Navarro, M. Coué, A. Kiouseloglou, P. Noé, F. Fillot, V. Delaye, A. Persico, A. Roule, M. Bernard, C. Sabbione, D. Blachier, V. Sousa, L. Perniola, S. Maitrejean, A. Cabrini, G. Torelli, P. Zuliani, R. Annunziata, E. Palumbo, M. Borghi, G. Reimbold, B. De Salvo, Trade-off between SET and data retention performance thanks to innovative materials for phase-change memory. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 9–11, 2013; Published: 2014
19. M. Boniardi, D. Ielmini, I. Tortorelli, et al., Impact of Ge-Sb-Te compound engineering on the set operation performance in phase-change memories. Solid State Electron. **58**, 11–16 (2011)
20. G.T. Jeong, et al., Process technologies for the integration of high density phase change RAM. IEEE International Conference on Integrated Circuit Design and Technology (ICICDT) Location: Austin, TX Date: Mat 9–11, 2005; Published: 2005
21. T.Y. Lee, K.H. Kim, D.S. Suh, et al., Low thermal conductivity in $Ge_2Sb_2Te_5$-SiO_x for phase change memory devices. Appl. Phys. Lett. **94**, 243103 (2009)
22. S. Braga, N. Pashkov, L. Perniola, A. Fantini, A. Cabrini, G. Torelli, V. Sousa, B. De Salvo, G. Reimbold, Effects of alloy composition on multilevel operation in self-heating Phase Change Memories. IEEE International Memory Workshop (IMW) Location: Monterey, CA Date: MAY 22–25, 2011; Published: 2011
23. A. Glière, O. Cueto, J. Hazart, Coupling the Level Set Method with an electrothermal solver to simulate GST based PCM cells. Proc. SISPAD, 63–66 (2011)
24. S.W. Ryu, H.K. Lyeo, J.H. Lee, et al., SiÜ2 doped Ge2Sb2Te5 thin films with high thermal efficiency for applications in phase change random access memory. Nanotechnology **22**, 254005 (2011)
25. Z. Wu, G. Zhang, Y. Park, et al., Controlled recrystallisation for low current RESET programming characteristics of phase change memory with Ge-doped SbTe. Appl. Phys. Lett. **99**, 143505 (2011)
26. C.F. Chen, et al., Endurance improvement of Ge2Sb2Te5-based phase change memory. IEEE International Memory Workshop (IMW) Location: Monterey, CA Date: MAY 10–14, 2009; Published: 2009

27. H.Y. Cheng, et al., Atomic-level engineering of phase change material for novel fast switching and high endurance PCM for storage class memory applications. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 9–11, 2013; Published: 2014
28. J. Kalb et al., Kinetics of crystal nucleation in undercooled droplets of Sb- and Te-based alloys used for phase change recording. J. Appl. Phys. **98**, p054910 (2005)
29. J. Kalb et al., Calorimetric measurements of phase transformations in thin films of amorphous Te alloys used for optical data storage. J. Appl. Phys. **93**(5), 2389 (2003)
30. V.M. Fokin, N.S. Yuritsyn, E.D. Zanotto, Nucleation and crystallisation kinetics in silicate glasses: Theory and experiment, in *Nucleation Theory and Applications*, ed. by W. P. Jürn (Wiley, Schmelzer, 2005), pp. 74–125
31. D. Turnbull, Correlation of liquid-solid interfacial energies calculated from supercooling of small droplets. J. Chem. Phys. **18**, 769 (1950)
32. M. Wuttig, C. Steimer, Phase change materials: From material science to novel storage devices. Appl. Phys. A Mater. Sci. Process. **87**, 411–417 (2007)
33. L. Perniola, V. Sousa, A. Fantini, et al., Electrical behavior of Phase-Change memory cells based on GeTe. IEEE Electron Device Letters **31**(5), 488–490 (2010)
34. X. Bicquart, M. Krbal, A. Kolobov, et al., Effect of doping on global and local order in crystalline GeTe. Appl. Phys. Lett. **98**, 231907 (2011)
35. A. Fantini, V. Sousa, L. Perniola, et al., N-doped GeTe as performance booster for embedded phase-change memories. IEEE International Electron Devices Meeting (IEDM) Location: San Francisco, CA Date: DEC 6–8, 2010; Published: 2011
36. J.H. Coombs, A.P.J.M. Jongenelis, W. van Es- Spiekman, B.A.J. Jacobs, Laser-induced crystallization phenomena in GeTe-based alloys. II. Composition dependence of nucleation and growth. J. Appl. Phys. **78**(8), 4918–4928 (1995)
37. H. Seo, T.-H. Jeong, J.-W. Park, et al., Investigation of crystallization behavior of sputter-deposited nitrogen-doped amorphous Ge2Sb2Te5 thin films. Jpn. J. Appl. Phys. **39**, 745–751 (2000)
38. S.J. Ahn et al., Highly reliable 50nm contact cell technology for 256Mb PRAM. Proc. IEEE VLSI **98**, 2201–2227 (2005)
39. Y.N. Hwang, et al., Writing current reduction for high density phase change RAM. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 8–10, 2003; Published: 2004
40. Q. Hubert, C. Jahan, V. Sousa, L. Perniola, A. Kusiak, J.-L. Battaglia, P. Noé, M. Bernard, C. Sabbione, M. Tessaire, F. Pierre, P. Zuliani, R. Annunziata, G. Pananakakis, B. de Salvo, A new insight on IRESET reduction of carbon-doped GST based PCM. Conference on Solid State Devices and Materials (SSDM) Location: Fukuoka, Japan (2013)
41. Q. Hubert, Optimisation de mémoires PCRAM pour générations sub-40 nm: intégration de matériaux alternatifs et structures innovantes. PhD dissertation, Univertity of Grenoble Alpes (2013)
42. L. Perniola, P. Noe, Q. Hubert, S. Souiki, G. Ghezzi, G. Navarro, A. Cabrini, A. Persico, V. Delaye, D. Blachier, J-P. Barnes, E. Henaff, M. Tessaire, E. Souchier, A. Roule, F. Fillot, J. Ferrand, A. Fargeix, F. Hippert, J-Y. Raty, C. Jahan, V. Sousa, G. Torelli, S. Maitrejean, B. De Salvo, G. Reimbold, Ti Impact in C-doped phase-change memories compliant to Pb-free soldering reflow. IEEE International Electron Devices Meeting (IEDM) Location: San Francisco, CA Date: DEC 10–13, 2012; Published: 2013
43. K.N. Chen et al., Thermal stress effects of $Ge_2Sb_2Te_5$ phase change material: Irreversible modification with Ti adhesion layers and segregation of Te. Microel. Eng. **85**, 2346–2349 (2008)
44. N. Matsuzaki, et al., Oxygen-doped GeSbTe phase change memory cells featuring 1.5V/100pA standard 0.13μm CMOS operations. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 5–7, 2005; Published: 2006
45. W. Czubatyj, et al., Current reduction in ovonic memory devices. European Phase Change and Ovonic Symposium (ePCOS) Location: Grenoble, France (2006)

46. W. Czubatyj et al., Nanocomposite phase change memory alloys for very high temperature data retention. IEEE Electron Device Lett **31**(5), 869–871 (2010)
47. H.Y. Cheng, T.H. Hsu, S. Raoux, et al., A high performance phase change memory with fast switching speed and high temperature retention by Engineering the GexSbyTez Phase Change Material. IEEE International Electron Devices Meeting (IEDM) Location: San Francisco, CA Date: DEC 10–13, 2012; Published: 2013
48. H.Y. Cheng, J.Y. Wu, R. Cheek, et al., A thermally robust Phase Change Memory by engineering Ge/N concentration in $(Ge, N)_x$ Sb_y Te_z phase change material. IEEE International Electron Devices Meeting (IEDM) Location: San Francisco, CA Date: DEC 10–13, 2012; Published: 2013
49. P. Zuliani, E. Varesi, E. Palumbo, et al., Overcoming temperature limitations in phase change memories with optimized GexSbyTez. IEEE Trans. Electron Devices **60**(12), 4020 (2013)
50. V. Sousa, G. Navarro, N. Castellani et al. Operation fundamentals in 12Mb Phase Change Memory based on innovative Ge-rich GST materials featuring high reliability performance. IEEE Symposium on VLSI Technology (VLSI Technology) Location : Kyoto, Japan Date: June 16-18, 2015; Published: 2015
51. S. Bordas, M.T. Clavaguera-Mora, Phase diagram of the ternary system Ge-Sb-Te. II. The subternary Ge-GeTe-Sb2Te3-Sb. Thermochimica Acta **107**, 239–265 (1986)
52. N. Ciocchini, E. Palumbo, M. Borghi, et al., Unified reliability modeling of Ge-rich phase change memory for embedded applications. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 9–11, 2013; Published: 2014
53. D. Ielmini, A. Lacaita, D. Mantegazza, Recovery and drift dynamics of resistance and threshold voltages in phase-change memories. IEEE Transactions on Electron Devices **54** (2), 308–315 (2007)
54. O. Cueto, V. Sousa, G. Navarro, et al., Coupling the phase-field method with an electrothermal solver to simulate phase change mechanisms in PCRAM Cells. IEEE International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) Location: Washington, DC Date: SEP 9–11, 2015; Published: 2015
55. G. Novielli, A. Ghetti, E. Varesi, et al., Atomic migration in phase change materials. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 9–11, 2013; Published: 2014
56. P. Zuliani, E. Palumbo, M. Borghi, et al., Engineering of chalcogenide materials for embedded applications of phase change memory. Solid State Electron. **111**, 27–31 (2015)
57. M.H.R. Lankhorst et al., Low-cost and nanoscale non-volatile memory concept for future silicon chips. Nat. Mater. **4**, 347–252 (2005)
58. Y.C. Chen, et al., Ultra-thin phase change bridge memory device using GeSb. IEEE International Electron Devices Meeting (IEDM) Location: San Francisco, CA Date: DEC 11–13, 2006; Published: 2007
59. I.S. Kim, et al., High performance PRAM cell scalable to sub-20nm technology below 4F2 cell size, extendable to DRAM applications. IEEE Symposium on VLSI Technology (VLSIT) Location: Honolulu, HI Date: June 15–17, 2010; Published: 2010
60. H.Y. Cheng et al., Crystallisation properties of materials along he pseudo-binary line between GeTe and Sb. JAP **115**, 093101 (2014)
61. M. Boniardi et al., Phase change memory: Device scaling and challenges for material engineering in the GeSbTe compound system. Microelectron. Eng. **137**, 1–4 (2015)
62. M. Boniardi, A. Redaelli, I. Tortorelli, et al., Electrical and thermal behavior of Tellurium poor GeSbTe compounds for phase change memory. IEEE International Memory Workshop (IMW) Location: Milan, Italy Date: MAY 20–23, 2012; Published: 2012
63. M. Micoulaut et al., Understanding amorphous phase-change materials from the viewpoint of Maxwell rigidity. Phys. Rev. B **81**, 174206 (2010)
64. H.Y. Cheng, W.C. Chien, M. BrightSky, et al., Novel fast switching and high data retention phase change memory based on new Ga-Sb-Ge material. IEEE International Electron Devices Meeting (IEDM) Location: Washington, DC Date: DEC 7–9, 2015; Published: 2016

Chapter 8
The Scaling of Phase-Change Memory Materials and Devices

Hasan Hayat, Krisztian I. Kohary, and C. David Wright

8.1 Introduction

In 1965 Gordon Moore of the Intel Corporation predicted that the number of transistors on an integrated circuit would double approximately every 1.5–2 years [1], so establishing the now well-known Moore's 'Law' and setting a target for the scaling of silicon-based electronic devices that designers still essentially follow today. The scaling of Si devices over the years is encapsulated by the historic trend in the so-called technology node, F, which has units of distance and represents the typical device half-pitch, i.e. half the distance between identical features in a device array (see Fig. 8.1). The smallest device size at any particular technology node is thus $4F^2$, and in Moore's day, since F was the order of several micrometres, the smallest devices occupied hundreds of square microns. By 2016, however, most state-of-the-art Si devices in production were being manufactured at the 22 or 14 nm nodes, giving typical device areas of around 1900 and 800 nm^2, respectively. A move towards the 10 and 7 nm nodes around 2017–2018 is expected for high-end chips (so device areas of down to 200 nm^2).

The scaling of devices generally brings with it a number of performance benefits, in particular increased speed and lower power consumption, along with, ultimately, a lower cost per bit to the end-user. This is true of the processor technologies for which Moore's Law was originally coined but also for established memory technologies such as silicon-based 'Flash' and DRAM (dynamic random access) devices; it is also true, as we shall see in this chapter, of phase-change memories (PCMs). Since, as we have already seen, silicon-based technologies can already be scaled down to around 200 nm^2 in terms of device area, any alternative memory technology should be capable of operating at significantly smaller sizes still. In this

H. Hayat • K.I. Kohary • C. David Wright (✉)
Department of Engineering, University of Exeter, Exeter EX4 4QF, UK
e-mail: David.Wright@exeter.ac.uk

© Springer International Publishing AG 2018
A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_8

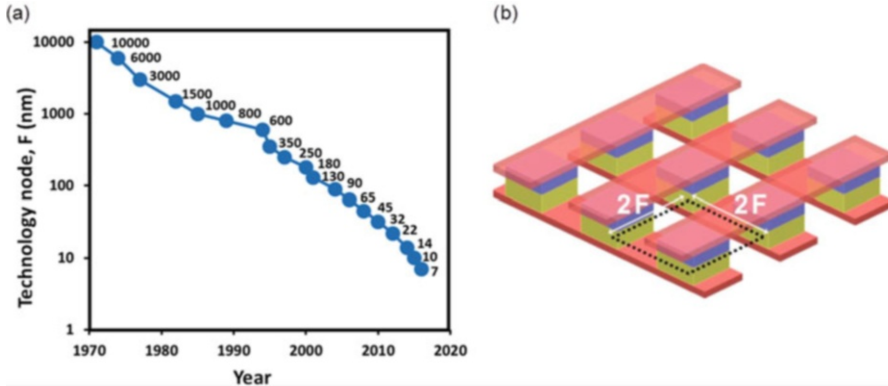


Fig. 8.1 (**a**) Historic trend in the semiconductor device technology node F (numbers on the plot refer to F value in nanometres). (**b**) Schematic of a generic device layout showing device pitch of F and smallest device area at node F equal to $4F^2$ (Reprinted with permission from [3]; copyright AIP Publishing LLC 2010)

chapter we will examine whether phase-change memories do indeed have such a capability.

Since phase-change memories rely on the repeated switching of the active region of a cell between the amorphous and crystalline states, and the ability on readout to differentiate electrically between these switched states, a number of fundamental questions naturally arise when we consider PCM scalability. For example, is there a minimum size (material volume) below which phase transformations in phase-change materials cease to be possible? Do reduced dimensions affect key material parameters, such as crystallization and melting temperatures, crystal nucleation rates and crystal growth velocities and thermal and electrical conductivity? And how does the device design itself affect the achievable scalability?

We might also ask about scalability in dimensions other than the spatial. For example, what controls, and how far can we scale down, the power consumption of phase-change devices? It is well known that the power limiting stage as far as PCM devices is concerned is the amorphization (or RESET) process, since this (usually) involves the heating of (at least a portion of) the phase-change material to above its melting temperature, which can be quite high (e.g. around 620 °C in $Ge_2Sb_2Te_5$). So, what material and device developments help us to reduce the RESET current? Similarly, we should consider the scaling of switching speed. What controls the switching speed of PCMs, and how fast can we make them?

We will address such issues in this chapter and try to answer the question of how far, and by what means, we can scale down the size, speed and power consumption of phase-change memory materials and devices while still retaining acceptable memory performance, by which we mean an acceptable read window (i.e. the difference in resistance between stored states), an acceptable device endurance (i.e. the number of switching cycles before device failure) and an acceptable retention time (i.e. the length of time a state may be stored and still successfully recalled) [2].

8.2 Scaling of Phase-Change Materials

One of the most important aspects for the development of PCM technology involves the memory material itself. Most technologically useful phase-change materials are formed from binary, ternary or quaternary chalcogenide alloys, i.e. alloys containing one or more elements from group V1A (group 16) of the periodic table, in particular tellurium (see Fig. 8.2). Some of the most notable phase-change materials used in applications lie on the pseudo-binary line in the

Fig. 8.2 (**a**) A map of Te-based phase-change materials plotted as a function of material ionicity and hybridization (Reprinted with permission from [9]; copyright Nature Publishing Group 2008). (**b**) The ternary Ge-Sb-Te phase diagram with some popular phase-change alloys highlighted (the red arrow indicates the trend of adding Ge to the so-called 'golden' composition $Ge_2Sb_1Te_2$ alloy)

Ge-Te-Sb ternary phase diagram (Fig. 8.2b) between the binary alloys Sb_2Te_3 and GeTe [4]. Thus, they have compositions of the form $(GeTe)_m(Sb_2Te_3)_n$, with the most well known being $(GeTe)_2(Sb_2Te_3)_1$ or $Ge_2Sb_2Te_5$ (commonly referred to as simply GST or GST-225) that has been used extensively in various rewritable optical disc formats (e.g. DVD-RAM) and is also somewhat of a 'standard' composition for PCM applications and against which all other compositions are compared/measured. Many other compositions, however, have also been explored and used successfully, for example, the combined doping with Ag and In of the binary eutectic alloy $Sb_{69}Te_{31}$ yielded the familiar Ag-In-Sb-Te (or AIST) alloy that was also used in rewritable optical discs (specifically the DVD-RW format). Other examples include the binary GeTe alloys [5], along with nitrogen- and carbon-doped variations of $Ge_2Sb_2Te_5$ for increased PCM thermal stability (see, e.g. [6, 7]). The so-called 'golden' composition of $Ge_2Sb_1Te_2$ (GST-212) is also gaining traction since it can offer a better overall compromise of material properties when compared to GST-225 [8]. Other notable material developments include the so-called phase-change superlattice materials that offer the intriguing possibility of non-melting phase transitions and so reduced switching powers (see Sect. 8.4.2).

The reason for exploring such a variety of compositions is the continual search for the 'perfect' phase-change material that can offer the best overall performance via a simultaneous combination of fast switching speeds (effectively fast crystallization), low-power operation (essentially low amorphization current), good thermal stability (a high crystallization temperature for high-temperature operation and extended data lifetimes) and excellent endurance (the number of switching cycles before failure). Such performance metrics are intrinsically linked to basic material properties including crystallization and melting temperatures, crystallization speeds, activation energies (for phase transformations) and electrical and thermal conductivities. When we turn our attention to the scaling possibilities and/or limitations for PCM devices, we must consider if and how such material properties vary as a function of size. This we do in the following sections, where we discuss the scaling behaviour of phase-change material properties in one, two and three dimensions.

8.2.1 Material Scaling in One Dimension

Scaling in one dimension consists essentially of reducing the thickness of the phase-change film (while keeping other dimensions constant). The most common observation when reducing film thickness is that the crystallization temperature increases. For example, Raoux and co-workers [10] used time-resolved X-ray diffraction (XRD) to study the crystallization behaviour of ultra-thin phase-change films with thicknesses in the range of 1–50 nm. Various materials were investigated, including $Ge_2Sb_2Te_5$ (GST), nitrogen-doped $Ge_2Sb_2Te_5$ (N-GST), $Ge_{15}Sb_{85}$, Sb_2Te and Ag- and In-doped Sb_2Te (AIST), with each film being sandwiched between Al_2O_3 layers for oxidation protection. In all cases, films with phase-

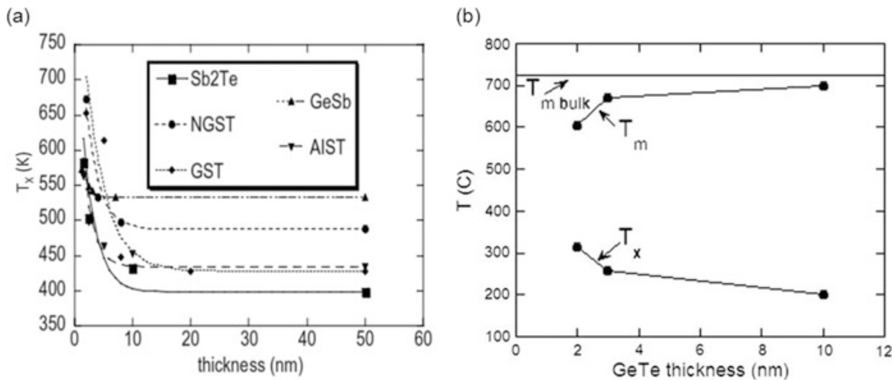


Fig. 8.3 (**a**) Variation of crystallization temperature as a function of film thickness for various phase-change materials (Reprinted with permission from [10]; copyright AIP Publishing LLC 2008). (**b**) Variation of crystallization and melting temperatures (T_x and T_m) versus film thickness for GeTe films (Reprinted with permission from [15]; copyright Elsevier 2008)

change layers having thicknesses between 10 and 50 nm showed little variation of crystallization temperature with thickness, but below 10 nm the crystallization temperature increased in some cases (i.e. for some materials) by as much as 200 °C (see Fig. 8.3). Films as thin as 2 nm for GST and N-GST, 1.5 nm for Sb_2Te and AgIn-Sb_2Te and 1.3 nm for GeSb were successfully crystallized, results that imply scaling (in 1D) of phase-change materials down to such ultra-thin dimensions should be possible while still ensuring reversible memory switching in device applications.

Raoux et al. found that the variation of crystallization temperature, T_x, with thickness agrees quite well with the empirical equation

$$T_x = T_{ax} + (T_m - T_{ax})e^{-d/C} \tag{8.1}$$

where T_{ax} is the crystallization temperature of the bulk material or a thick film, T_m is the melting temperature, d is the film thickness and C is a fitting constant. This empirical equation can be related to the Gibbs free energy of the system via

$$T_m = T_{ax}\left(\frac{S_{oc} - S_{oa}}{S_{ac}}\right) \tag{8.2}$$

where S_{oc} is the free (surface) energy of an oxide-crystalline interface and S_{oa} and S_{ac} the equivalent surface energy terms for oxide-amorphous and amorphous-crystalline interfaces. Thus, the change in crystallization temperature with thickness is determined by the interfacial energies, the melting temperature and the fitting parameter C (which is related to the average screening or bonding length typical for the oxide encapsulating layer and the phase-change layer in its crystalline form – see [10]).

It is clear from Eq. 8.1 that the encapsulation layer(s) are likely to play a significant role in the crystallization behaviour of thin phase-change films. Indeed, Simpson et al. [11] found that while TiN capping layers (TiN is a common electrode material in PCM devices) produced similar increases in crystallization temperature when film thickness was reduced as seen for the Al_2O_3 layers used by Raoux, the use of $(ZnS)_{0.85}(SiO)_{0.15}$ capping layers resulted in no increase (even a small decrease) in crystallization temperature as GST(−225) films were reduced in thickness from 10 to 2 nm. Such differences were attributed to the very different levels of compressive stress induced in the GST layer by the different capping layers (58 and 240 MPa for $(ZnS)_{0.85}(SiO)_{0.15}$ and TiN, respectively). Thus, the increase in crystallization temperature with reducing film thickness shown in Fig. 8.3 appears not to be a universal observation for all encapsulating materials. Nonetheless, Simpson et al. also found that films just 2 nm thick were still able to crystallize, which, in terms of scaling, is the most pertinent finding.

While the variation of crystallization temperature, along with the ability to crystallize at all, as film thickness is scaled down is a key issue, we should also consider the variation of other important material parameters, such as activation energy, crystallization speed and melting temperature. Activation energies are typically accessed via measurements of the crystallization temperature versus heating rate followed by a Kissinger analysis (i.e. fitting of the results to the Kissinger equation). Since in the majority of cases reported the crystallization temperature was shown to increase for thinner phase-change layers, it is not a surprise that the activation energy (E_A) for crystallization is also in general found to increase with reducing thickness (since the two are related in the Kissinger equation). For example, $Ge_2Sb_2Te_5$ showed an increase of E_A from 2.86 eV to 4.66 eV as film thickness was reduced from 20 to 5 nm [12].

Turning our attention to crystallization speed, here we have to consider both the crystal nucleation rate (i.e. the number/unit volume/unit time of crystal nuclei that form) and the crystal growth rate (i.e. the distance/unit time by which an existing crystallite grows). Some materials, such as GST-225, are nucleation dominated (i.e. tend to crystallize by the formation of many nuclei that grow and merge), while others, such as AIST, are growth dominated (tend to crystallize by the formation of relatively few nuclei that grow rapidly), although in reality both nucleation and growth can happen simultaneously, and both should be considered in most materials. Since the crystallization process is also significantly affected by the type of capping/encapsulation layers used (see previous paragraphs), it is also no surprise that sometimes conflicting observations of the effects of film thickness on crystallization speed have been reported. For example, the crystallization speed has been observed to decrease for $Ge_2Sb_2Te_5$ (sandwiched between ZnS-SiO_2 dielectric layers) as film thickness reduces, whereas for AIST it increased [13]. A fast-growth doped SbTe alloy composition, on the other hand, showed an optimum thickness (of 9 nm) at which crystallization speed was fastest [14]. A general observation then is that the crystallization speed can either increase (which is desirable) or decrease (undesirable) as film thicknesses are reduced, depending on the phase-change material composition and capping layers.

The melting temperature T_m also plays a key role in the operation of PCM devices, since (at least with conventional materials) the phase-change layer has to be heated to above T_m in order to melt- quench it into the amorphous phase. Melting temperatures differ considerably for different phase-change alloys but are usually in the range of 400–800 °C. There are considerably fewer studies of the effect of film thickness on melting temperature *cf*. the effect on crystallization temperature, with the most common finding being a reduction in T_m as film thickness decreases. For example, Raoux et al. [15] found that the melting temperature of GeTe thin films decreased as the film thickness was reduced from 10 to 2 nm (see Fig. 8.3b), while Her and Hsu [16] reported a drop in T_m for $Sb_{70}Te_{30}$ alloys from 428 to 387 °C as film thickness reduced from 70 to 10 nm.

Finally, in this section, we consider the effects of film thickness on thermal conductivity, *k*, since the thermal conductivity in devices affects a number of important operating parameters, such as input power required (to reach a given temperature), heating and cooling rates, etc. In general, the thermal conductivity of phase-change films encapsulated by dielectric and/or metal layers shows a decrease as film thickness decreases. For example, Reifenberg et al. [17] reported *k* values of 0.29, 0.42 and 1.76 $Wm^{-1}K^{-1}$ for as-deposited amorphous, fcc and hcp phases of 350 nm $Ge_2Sb_2Te_5$ films reducing to 0.17, 0.28 and 0.83 $Wm^{-1}K^{-1}$, respectively, as the film thickness was reduced to 60 nm. However, measured thermal conductivity values are invariably affected by thermal boundary resistances (TBR, i.e. the thermal resistance between the phase-change layer and any encapsulating or electrode layers), with TBR effects becoming more important as the film thickness is reduced, making it important to decouple properly intrinsic conductivity and boundary resistance effects as films get thinner and thinner [18, 19].

8.2.2 Material Scaling in Two Dimensions

Material scaling in two dimensions has been studied via the fabrication and characterization of phase-change nanowires. Such nanowires are typically synthesized using the metal catalyst-mediated vapour-liquid-solid (VLS) processes, have diameters ranging from a few tens to hundreds of nanometres and are usually single crystal in their as-deposited state (see Fig. 8.4) [20–22].

Single-crystal GeTe nanowires (40–80 nm in diameter) have been shown to melt at a temperature of 390 °C as compared to bulk GeTe (725 °C) [20]. A similar effect was observed for 40–80 nm diameter In_2Se_3 nanowires, where the melting temperature dropped from 890 °C (for bulk) to 690 °C for the smallest nanowires [23]. The reduction of melting temperature with size is advantageous, as a reduced melting point will consequently decrease the current and power required to amorphize the phase-change material (during the so-called RESET process). Indeed, such a reduction in power consumption was noted by Lee et al., for $Ge_2Sb_2Te_5$ nanowires, where shrinking the nanowire diameter from 200 to 30 nm resulted in a power reduction factor for amorphization of over 3.5 (see Fig. 8.4b and [21]). Lee et al.

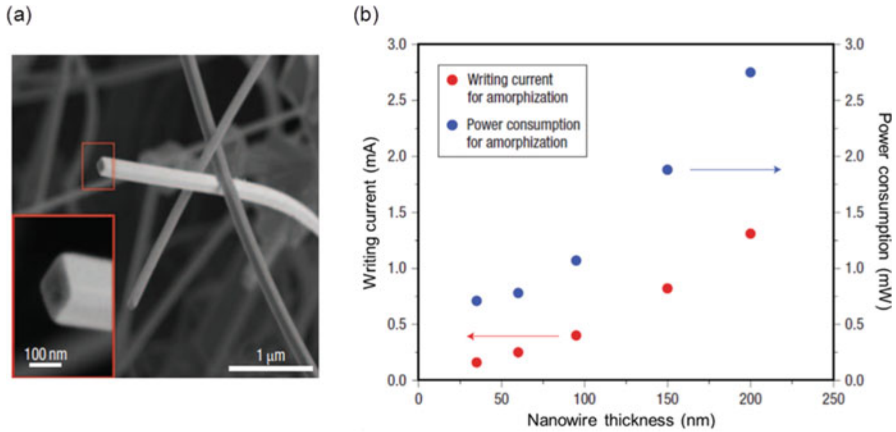


Fig. 8.4 (**a**) SEM image of as-grown $Ge_2Sb_2Te_5$ nanowires. (**b**) The variation in amorphization current and power as a function of $Ge_2Sb_2Te_5$ nanowire diameter (Reprinted with permission from [21]; copyright Nature Publishing Group 2007)

also showed that $Ge_2Sb_2Te_5$ nanowires could be repeatedly switched between amorphous and crystalline states (over 10^5 times) and that they had very long predicted data retention times, at least for larger diameters (e.g. 1800 years at 80 °C for 200 nm wires). Interestingly, in contrast to typical findings for thin-film phase-change layers, the activation energy (for crystallization) of the $Ge_2Sb_2Te_5$ nanowires decreased as the nanowire diameter decreased (for thin films activation energy typically increases as films get thinner), a behaviour attributed to an increased surface-to-volume ratio and heterogeneous nucleation effects in the nanowires.

The most interesting observation in phase-change nanowires is the phenomenon of non-melting amorphization; while this is not necessarily a size-scaling phenomenon, the prospect of a non-melting amorphization is particularly attractive for PCM applications since it promises significant reductions in power consumption. Nam et al. demonstrated such a non-melting crystal-to-amorphous transformation process in single-crystal $Ge_2Sb_2Te_5$ nanowires, a process explained in terms of the formation of disorder via the build-up of electrically driven dislocations [24]. The same research group also showed that by the control of defect formation using ion implantation, non-melt-quenched amorphization in GeTe nanowires could be successfully induced, resulting in very significant reductions in the amorphization (RESET) current and current densities (and hence powers) as compared to a melt-quenched approach. For example, ion-implanted GeTe nanowires of size 80 × 80 nm^2 showed RESET currents and current densities of only 8 μA and 0.13 $MAcm^{-2}$, respectively, compared to currents of many hundreds of μA and current density of at least 50 $MAcm^{-2}$ for notionally identical nanowires amorphized by melt quenching (see Fig. 8.5 and [25]).

Nanowire-type size scaling has been taken even further by Giusca et al. who used carbon nanotubes (CNTs) as templates, growing ultrasmall GeTe nanowires of

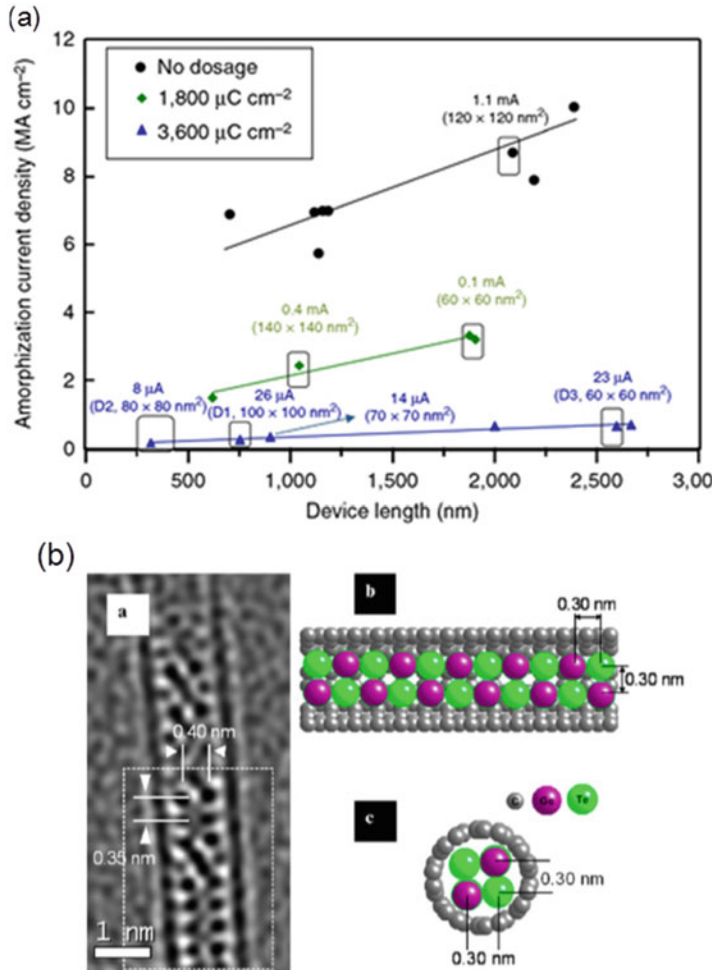


Fig. 8.5 (**a**) Amorphization current densities in GeTe nanowires; circles show conventional nanowires having a melt-quenched amorphization process, and diamonds and triangles show results for nanowires defect-engineered to provide non-melt-quenched amorphization (numerical values adjacent to specific data points show nanowire device contact area and reset current amplitude) (Reprinted with permission from [25]; copyright Nature Publishing Group 2016). (**b**) GeTe nanowires grown inside a 1.3 nm CNT template (Reprinted with permission from [26]; copyright American Chemical Society 2013)

diameters < 2 nm inside (see Fig. 8.5b [26]). Such templated nanowires were found to crystallize and display amorphous-to-crystalline phase changes on these sub-2 nm sizes, demonstrating quite clearly the promising potential for phase-change devices to be scaled to ultrasmall dimensions.

8.2.3 Material Scaling in Three Dimensions

Scaling in three dimensions has been studied via the deposition and characterization of phase-change nanoparticles and nanoclusters of various types (see, e.g. Fig. 8.6). For example, $Ge_2Sb_2Te_5$ nanoparticles of sizes ranging from 4 to 50 nm in diameter were fabricated in several studies via laser ablation techniques [27–30], while the fabrication via e-beam lithography of $Ge_2Sb_2Te_5$, nitrogen-doped $Ge_2Sb_2Te_5$, $Ge_{15}Sb_{85}$, Sb_2Te and AIST nanoparticles (in the size range 20–80 nm, encapsulated by 8 nm Al_2O_3) has been reported by Raoux et al. [31]. X-ray diffraction (XRD) measurements by Raoux et al. showed that all the fabricated nanoparticle arrays crystallized at temperatures similar to those of similarly prepared blanket films,

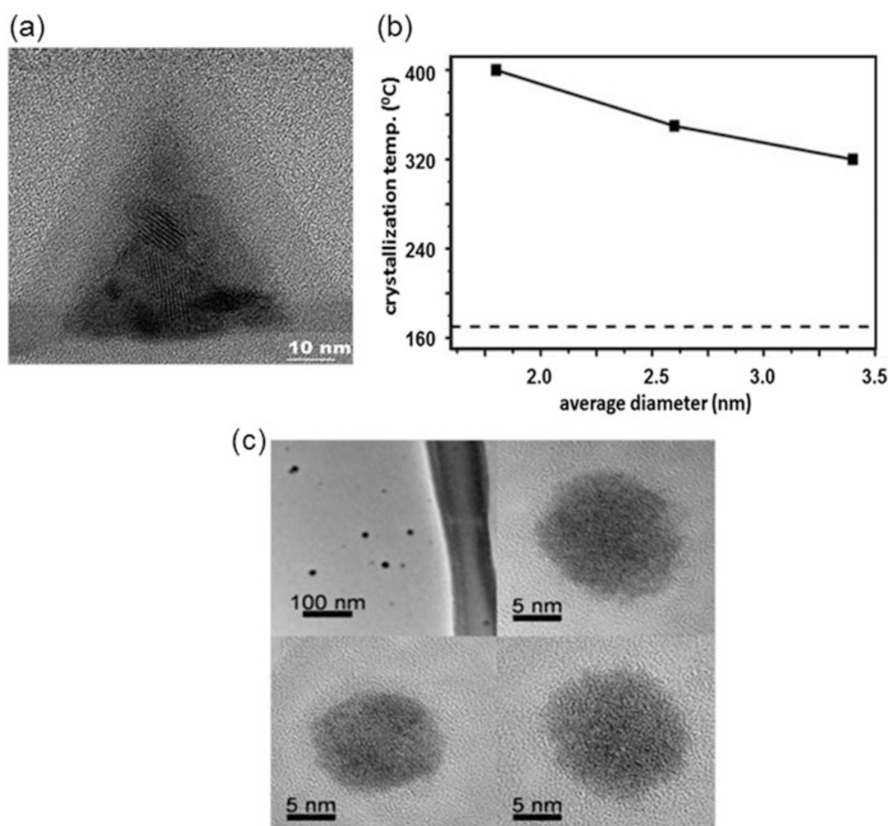


Fig. 8.6 (**a**) TEM image of an annealed (crystallized) GeSb nanoparticle (Reprinted with permission from [31]; copyright AIP Publishing LLC 2007). (**b**) Variation of crystallization temperature with nanoparticle diameter for ultrasmall GeTe nanoparticles (Reprinted with permission from [33]; copyright Royal Society of Chemistry 2009). (**c**) TEM images of as-deposited amorphous $Ge_2Sb_2Te_5$ nanoclusters (Image provided courtesy of Giada Ghezzi (CEA Leti) and Robert Morel (CEA Inac) and taken from [35])

except for the case of Sb_2Te which crystallized at a temperature 40 °C higher than its blanket film counterpart. However, on shrinking the nanoparticles still further to 15 nm and below (in subsequent studies) using a self-assembly approach [32], the same group found that crystallization temperatures increased as compared to the bulk values (in line with results found for ultra-thin films, see Sect. 8.2.1). Such increases in crystallization temperatures (cf. bulk values) with shrinking nanoparticle size were also seen by Caldwell et al. [33] for GeTe, where nanoparticles as small as 1.8 nm were observed to have a crystallization temperature more than double that of the reported bulk value (see Fig. 8.6b). Caldwell et al. predicted using these results that somewhere not much below 1.8 nm is, approximately, the ultimate scaling limit for GeTe phase-change materials, since at such ultrasmall sizes, the crystallization and melting temperatures will coincide, and the material will never crystallize but will rather transition directly from the amorphous phase to the melted phase [33].

While many of the nanoparticles discussed above were prepared using some kind of lithographic or self-assembly approach, Ghezzi et al. reported the direct gas-phase sputter (condensation) deposition of $Ge_2Sb_2Te_5$ nanoclusters [34]. Such nanoclusters are almost spherical in shape, had a diameter of ~6 nm and were embedded within a layer of alumina, yielding well-defined, contaminant-free and isolated clusters. XRD studies showed that the as-grown clusters were amorphous in phase but transformed on heating to the cubic fcc phase (with a crystallization temperature of ~180 °C, some 25 °C higher than a 10 nm thin film of the same composition), demonstrating clearly the potential for phase-change switching in volumes as small as ~100 nm^3.

The experimental results discussed above point to minimum nanoparticle sizes of around 1–2 nm in order to achieve proper crystallization in a wide range of phase-change materials (with the actual minimum value dependent on the actual material used, the type of encapsulation layer, substrate type, etc.). Such a value is very similar to the stable crystallite cluster size, n_c, obtained via classical nucleation theory and resulting from a balancing of volume and surface energies present when a crystallite grows in an amorphous matrix and is known to occur when (for spherical crystallites)

$$n_c = \frac{32\pi {\nu_m}^2 S^3}{3\Delta\gamma^3} \tag{8.3}$$

where v_m is the volume of the crystalline 'species' (e.g. a 'monomer' of $Ge_2Sb_2Te_5$, GeTe, etc.), S is the interfacial surface energy density between the amorphous and crystalline phases and $\Delta\gamma$ is the bulk free energy difference per 'monomer'. The parameters in the above equation have been previously estimated for $Ge_2Sb_2Te_5$ [36] and lead to a variation of the minimum stable crystallite size as a function of temperature as shown in Fig. 8.7, where it can be seen that (spherical) crystallites of around 1–2 nm are predicted to be stable for temperatures up to around 400 °C [37].

Interestingly, atomistic molecular dynamics simulations based on density functional theory have also predicted the critical crystal nucleus size in $Ge_2Sb_2Te_5$ to be

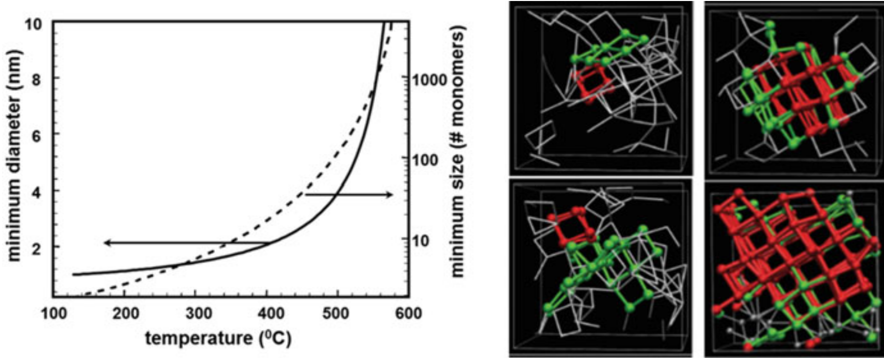


Fig. 8.7 (**a**) Predicted (via classical nucleation theory) minimum stable crystallite cluster size for $Ge_2Sb_2Te_5$ as a function of temperature (Reprinted with permission from [37]; copyright IEEE 2006). (**b**) Atomic configurations during the crystallization process in amorphous $Ge_2Sb_2Te_5$ and the evolution of structural units on annealing at 600 K. (top left) Formation of structural units during the incubation period: a significant number of fourfold rings (silver) exist, but only a few planes (green) or cubes (red) are formed occasionally; (bottom left) development of ordered layer structures at the crystallization site; (top right) a cube cluster and planes extending from the cluster interface; (bottom right) completely crystallized phase with a crystal-glass interface (Reprinted with permission from [38]; copyright American Physical Society 2011)

of similar order to that predicted by the much simpler classical nucleation theory. Specifically Lee and Elliott [38] simulated the formation and growth of crystalline clusters (see Fig. 8.7b) in amorphous $Ge_2Sb_2Te_5$ upon thermal annealing (at 600 K or 337 °C) and found that once the size of clusters became larger than approximately five to ten connected $(GeSb)_4Te_4$ cubes (which would occupy a volume of around 1.5–3 nm^3), they successfully started to grow, rather than decay. Furthermore, Simpson et al. [11], using reasoning based on resonant bonding and the minimum number of vacancies needed to ensure crystallization of $Ge_2Sb_2Te_5$ in the cubic (fcc) phase, argued that the minimum volume required for memory operation (for $Ge_2Sb_2Te_5$) is 1.7 nm^3. Thus, there seems, at least for $Ge_2Sb_2Te_5$, unanimity via various theoretical approaches with regard to the smallest volume of material, at ~2 nm^3, that could provide phase-change memory functionality, and this value agrees well with experimental studies of the crystallization processes in phase-change nanoclusters.

8.3 Scaling of Phase-Change Memory Devices

From the experimental and theoretical studies discussed in Sects. 8.2.1, 8.2.2, 8.2.3, it is evident that phase-change materials have the capability to operate down to very small dimensions, indeed, down to volumes of only a few cubic nanometres. This is most promising in terms of the future development of PCM devices. However, the development of real devices on such size scales is not straightforward, in particular

if the goal is that of commercially viable device designs. In this section, therefore, we examine the design and development of phase-change memory devices, concentrating in particular on approaches that are used to decrease device size, decrease the switching current and power and increase the switching speed.

8.3.1 The PCM 'Mushroom' Cell

A typical PCM cell consists of the active phase-change material (e.g. GST) sandwiched between two metal electrodes (e.g. TiN and W) and insulated from adjacent cells using a dielectric material such as SiO_2. The most established structure in this regard is the so-called 'mushroom' (or 'lance') cell, for which a schematic and TEM image is shown in Fig. 8.8 and in which the phase-change material and top electrode are planar layers deposited and patterned on a pillar-like bottom 'heater' contact. Note that this bottom contact is usually fabricated from doped TiN and has a higher electrical resistance (and lower thermal conductivity) compared to usual metal electrode materials. In normal device operation, a portion of the phase-change layer sitting directly on top of the heater contact is switched between the amorphous and crystal phases by appropriate electrical excitations (RESET and SET pulses, respectively), such that the active region within the phase-change layer and the heater pillar resemble a mushroom in shape, hence the name (see the blue- and orange-shaded regions in Fig. 8.8a and the amorphized dome in TEM image of Fig. 8.8b).

The use of the pillar-type contact in the mushroom cell is attractive since, as is obvious from Fig. 8.8, it helps to limit the volume of phase-change material that has to be melted/amorphized and recrystallized, thus also limiting the RESET current and power and increasing the recrystallization speed. The obvious question that follows is how far, practically, can mushroom-type cells be scaled down in size, and what effect does such shrinking have on device performance?

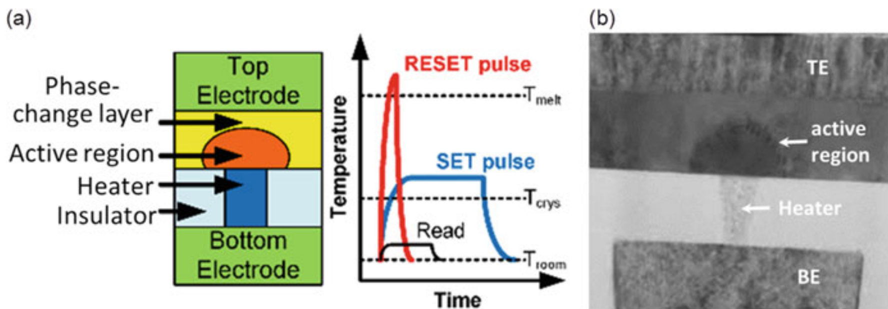


Fig. 8.8 (**a**) Schematic of a typical mushroom cell structure widely used for PCM devices (and schematic of write and read pules). (**b**) TEM cross-section of a mushroom-type PCM showing the rounded amorphous dome formed above the heater (Reprinted with permission from [39]; copyright AIP Publishing LLC 2011)

8.3.1.1 Analytical and Numerical Studies of Scaling in PCM Mushroom-Type Cells

One of the first scaling studies of mushroom-type PCM cells was reported by Pirovano et al. in 2003, using GST-225 for the phase-change layer [40] (see also Chap. 4). Assuming isotropic scaling (i.e. assuming that the heater contact diameter, the height of the heater pillar and the thickness of the phase-change layer all scale linearly by a factor k (where $k > 1$) as the device shrinks), then key device parameters were predicted to scale as shown in Table 8.1.

The scaling behaviour shown in Table 8.1 can be explained by a simple analytical model (see also discussion in Chap. 4) for the programming (RESET) of a PCM mushroom cell, as expounded by Russo et al. [41]. The model is based on an ohmic representation of the cell, as shown in Fig. 8.9. It assumes that the main contribution to overall heating occurs within a (hot) region of diameter D close to the phase-change layer/heater interface. In this framework, the temperature T in the hot region during the current pulse can be calculated as

Table 8.1 Scaling of PCM device operating parameters as the device is scaled isotropically in size by the factor k (for $k > 1$)

Device parameter	Scaling factor
Heater contact diameter	$1/k$
Heater contact area	$1/k^2$
Phase-change layer thickness	$1/k$
Heater height	$1/k$
Thermal resistance	k
SET resistance	k
RESET resistance	k
ON-state resistance	k
Programming (RESET) current	$1/k$
Programming voltage	1
Programming power dissipation	$1/k$

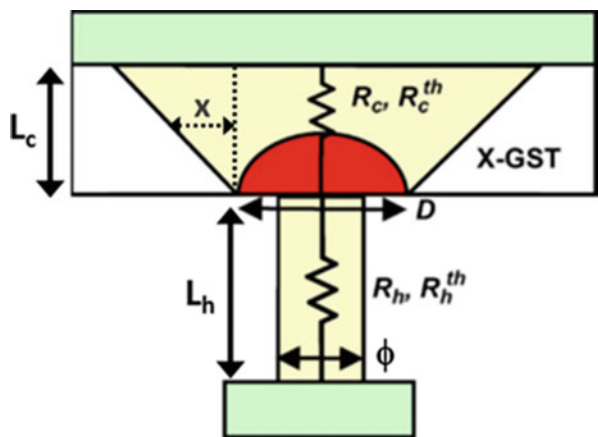


Fig. 8.9 Schematic of a PCM cell showing key geometrical and physical parameters used in the analytical scaling study (Reprinted with permission from [41]; copyright IEEE 2008)

$$T = T_0 + \beta P_{\mathrm{GEN}} R^{\mathrm{th}} \tag{8.4}$$

where T_0 is ambient (room) temperature, P_{GEN} is the dissipated power within the cell, R^{th} is the overall thermal resistance from the hot spot to the heat sinks (at T) and β is the fraction of total power dissipated in the hot region. P_{GEN} in turn is given approximately by

$$P_{\mathrm{GEN}} = R_{\mathrm{ON}} I^2 \tag{8.5}$$

where R_{ON} is the (electrical) resistance of the cell during programming. R_{ON} is given by the serial combination of the heater resistance, R_{h}, and the phase-change layer resistance, R_{c}, i.e.

$$R_{\mathrm{ON}} = R_{\mathrm{h}} + R_{\mathrm{c}} \tag{8.6}$$

The thermal resistance, R^{th}, on the other hand, is given by the parallel combination of the thermal resistance of the phase-change layer, $R_{\mathrm{c}}^{\mathrm{th}}$, and that of the heater, $R_{\mathrm{h}}^{\mathrm{th}}$ (since heat can flow towards both top and bottom electrodes simultaneously), i.e.

$$\frac{1}{R^{\mathrm{th}}} = \frac{1}{R_{\mathrm{h}}^{\mathrm{th}}} + \frac{1}{R_{\mathrm{c}}^{\mathrm{th}}} \tag{8.7}$$

The melting current I_{m} then follows from Eq.(8.5), substituting for P_{GEN} from Eq.(8.4) with $T = T_{\mathrm{m}}$ and R^{th} from Eq.(8.7) to yield

$$I_{\mathrm{m}} = \sqrt{\frac{T_{\mathrm{m}} - T_0}{\beta R_{\mathrm{ON}}} \left(\frac{1}{R_{\mathrm{h}}^{\mathrm{th}} + R_{\mathrm{c}}^{\mathrm{th}}} \right)} \tag{8.8}$$

To determine how I_{m} (and thus the RESET current) scales with the scaling factor k, we need in turn to determine how $R_{\mathrm{h}}^{\mathrm{th}}$, $R_{\mathrm{c}}^{\mathrm{th}}$ and R_{ON} all scale with k. The heater thermal resistance can be written (by analogy with electrical resistance $R = \rho L/A$) simply as

$$R_{\mathrm{h}}^{\mathrm{th}} = \frac{1}{\theta_{\mathrm{h}}} \frac{L_{\mathrm{h}}}{\pi r_{\mathrm{h}}^2} \tag{8.9}$$

where θ_{h} is the heater thermal conductivity, r_{h} the heater radius and L_{h} its length. The thermal resistance presented by the phase-change layer can be found by calculating the heat flow towards the top electrode (the top thermal contact). By assuming the heat flows in a truncated cone of base diameter equal to that of the hot spot (D) and having an angle of 45° (see Fig. 8.9), then $R_{\mathrm{c}}^{\mathrm{th}}$ can be calculated from

$$R_{\mathrm{c}}^{\mathrm{th}} = 2 \int_0^{L_{\mathrm{c}}} \frac{1}{\theta_{\mathrm{c}}} \frac{1}{\pi (D/2 + x)^2} dx \tag{8.10}$$

where θ_c is the thermal conductivity of the phase-change layer and L_c its thickness. Solution of the integral in Eq.(8.10) yields

$$R_c^{th} = \frac{1}{\theta_c} \frac{4L_c}{\pi D(D + 2L_c)} \tag{8.11}$$

Similarly, the electrical resistance of the heater is simply

$$R_h = \rho_h \frac{L_h}{\pi r_h^2} \tag{8.12}$$

where ρ_h is the heater's electrical resistivity. The resistance of the phase-change layer can be given by analogy with Eq. (8.11), replacing D with the electrical contact diameter (i.e. the diameter of the heater, ϕ), such that

$$R_c = \rho_c \frac{4L_c}{\pi\phi\ (\phi + 2L_c)} \tag{8.13}$$

Inspecting Eqs. (8.9), (8.11), (8.12) and (8.13), we can see that as L_h, L_c and ϕ ($= 2r_h$) are all scaled isotropically as $1/k$ (and note that D scales isotropically too), then R_h^{th}, R_c^{th} and R_{ON} (where $R_{ON} = R_c + R_h$) will all scale as k. Thus, from Eq. (8.8) it is easy to that, under the conditions of isotropic scaling of the PCM cell by a factor k, that the melting current I_m must scale as $1/k$, as shown in Table 8.1.

Similarly, since the melting voltage V_m is given simply by

$$V_m = I_m R_{ON} \tag{8.14}$$

and the power by

$$P_m = V_m I_m \tag{8.15}$$

we see that the melting voltage is constant with scaling (I_m scales as $1/k$, R_{ON} scales as k), while the power dissipation for melting scales as $1/k$ (I_m scales as $1/k$, V_m independent of k). The key role played by the scaling of the contact area (and consequently the volume of phase-change material that undergoes melting) in reducing melting currents and powers (and hence RESET currents and powers) is thus clearly evident, even from a simple analytical approach such as that presented above.

Pirovano [40] and Russo [41] used analytical models such as those above, backed up by numerical electrical and thermal simulations, to predict the variations of key device operating parameters as device sizes shrink. Typical results, assuming a GST phase-change layer and a TiN heater, are shown in Fig. 8.10 (taken from [41]) where the values of R_{ON} and I_m are shown as a function of the technology node, F, down to the 16 nm node (note that in Fig. 8.10a, ϕ is the heater contact diameter ($= 2r_h$)). It is clear that shrinking of the device leads to improved performance in key areas, in particular the reduction of the melting (RESET)

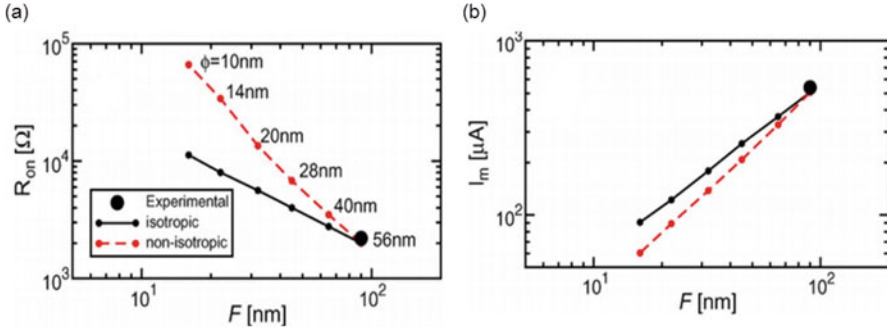


Fig. 8.10 (**a**) Variation of PCM mushroom-type cell ON resistance as a function of the technology node size (*F*); two cases are shown assuming isotropic and non-isotropic scaling (see text). Numerical values for ϕ refer to the heater contact diameter. (**b**) Variation in the melting (RESET) current as a function of the technology node (*F*) (Both reprinted with permission from [41]; copyright IEEE 2008)

current, which drops from over 500 μA at the 90 nm node (heater diameter of 56 nm or ~2500 nm^2 heater contact area) to under 100 μA at the 16 nm node (heater diameter of 10 nm or ~80 nm^2 heater contact area). The fact that the operating characteristics of PCM devices improve as the devices become smaller is one of the most attractive features of this technology – things get better as we make them smaller.

Of course isotropic device scaling of the mushroom-type cell is not the only approach to reducing critical operating parameters, such as the RESET current; there is also the possibility of cell geometry optimization, the use of alternative cell designs as well as the use of alternative materials. Let's consider first an approach based on possible optimizations of the mushroom cell geometry. In the isotropic scaling discussed previously, the phase-change layer thickness (L_c), the heater length (L_h) and the heater contact diameter ϕ were all scaled simultaneously by the same factor. However, this need not be the case. One form of non-isotropic scaling is to keep L_c and L_h constant while reducing the heater contact diameter ϕ (note that ϕ is the key lithographically limited dimension). Results for this case are also shown in Fig. 8.10, where it can be seen that additional benefits, in terms of a stronger reduction in the melting current with decreasing technology node, are possible. It should however be noted that we cannot allow L_c and L_h to take on arbitrary values just so as to reduce I_m. This is due to both manufacturing limitations and, importantly, the effect that L_c and L_h have on the resistance of the PCM cell (see Fig. 8.11), in particular the resistance, R_{SET}, when the cell is in the crystalline state. This is because the value of R_{SET} affects the size of the readout current; the higher R_{SET}, the smaller the readout current and the longer it takes to sense the state of the cell during readout [42]. Since the values chosen for L_c and L_h will also alter the thermal (in addition to electrical) resistances, as is evident from Eqs. (8.6) to (8.9), they also affect very considerably the temperature distribution in the cell, as

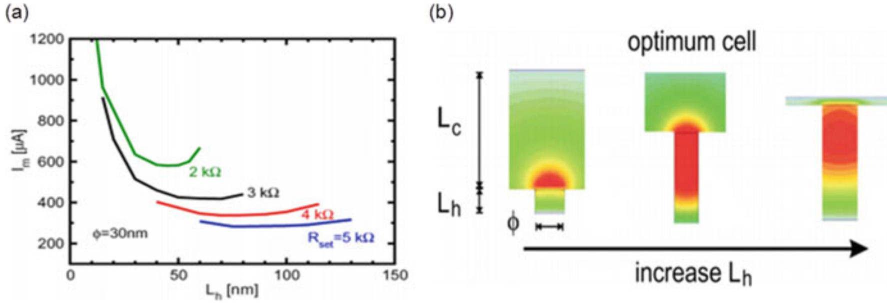


Fig. 8.11 (**a**) The variation of melting current, I_m, as a function of heater length, L_h, for a fixed heater diameter of 30 nm and for a range of cell SET resistances, R_{SET} (the thickness, L_C, of the phase-change layer, GST in this case, was varied to maintain a constant R_{SET} as L_h varied). (**b**) Simulated temperature distributions in mushroom-type PCM cells (each subject to their own individual melting current) with different L_c and L_h sizes (Both reprinted with permission from [41]; copyright IEEE 2008)

shown in Fig. 8.11b. Note that the 'optimum cell' shown in Fig. 8.11b (and where the hot spot is located close to the heater/phase-change layer interface) is the structure chosen for the isotropic scaling study, the results of which are shown in Fig. 8.10.

Another important consideration when device sizes shrink is the possible effects of thermal crosstalk, i.e. the thermal disturbance effects that programming of one cell might have on adjacent cells. Such 'parasitic' heating could lead to the partial recrystallization of an amorphous bit, building up over many cycles to the unwanted erasure of data. Fortunately it was shown by Russo et al. [43] that isotropically scaled PCM mushroom-type devices can be expected not to suffer from any significant thermal crosstalk problems, although non-isotropic scaling approaches could cause problems.

Despite the relative maturity of the mushroom-type PCM cell, it appears that the smallest of such cells to have been successfully fabricated to date (or at least the smallest that have been publicly reported) are at the 90 nm node, with a corresponding heater contact diameter of approximately 60 nm [44]. In similar vein, most scaling studies, such as those discussed above, have been either analytical in nature or have not addressed the likely performance of mushroom cells for sub-10 nm heater sizes. In an attempt to redress this omission, we ourselves have recently used a physically realistic computational model that combines electrical, thermal and phase transformation simulations to assess the likely performance of ultrasmall GST-based PCM mushroom cells [45]. The electrical and thermal models are implemented using finite element software (COMSOL™) and solve, simultaneously, the Laplace and heat diffusion equations to calculate at each time step and for any given electrical excitation the 3D (or pseudo-3D, i.e. 2D with assumed cylindrical symmetry in the case of the mushroom-type cell) temperature distribution throughout the cell. This temperature distribution drives a phase transformation model, for which we use a Gillespie Cellular Automata (GCA) approach

that combines thermodynamic features of rate-equation methods with elements from probabilistic cellular automata (PCA) and phase-field models and uses the Gillespie algorithm for efficient time-stepping. Our GCA model has been previously described in detail [46] but in summary considers a homogeneous, isotropic material in a square lattice where the state of the material is described through a set of points in the lattice that can be either crystalline or amorphous. The state of each point (i,j) in the lattice is described by two quantities: r_{ij}, the phase of the (i,j) site (which takes the values 0 and 1 for amorphous and crystalline, respectively), and Φ_{ij}, which defines an orientation (with two adjacent crystalline sites belonging to same crystallite (crystal grain) if they have the same orientation). The local changes that can occur are defined by three events: nucleation, where site (i,j) and an adjacent site, originally both amorphous, become a single crystallite; growth, where site (i,j), originally amorphous, becomes attached to an adjacent crystal; and dissociation, where site (i,j), originally crystalline, detaches from the crystal of which it is a part of to become amorphous. The rate at which each of these three events occurs is determined by the system energy, which is usually described in terms of the Gibbs free energy G, where $G = (AS - Vg)$ and A and V are the surface area and volume, respectively, of a crystal cluster, S is the surface energy and g the bulk free energy difference between phases. The bulk energy difference term g is considered to be purely temperature dependent (e.g. as $g(T) = H_f\,(7\,T/T_m)[(T_m - T)/(T_m + 6\,T)]$ where H_f is the enthalpy of fusion and T_m is the melting point [36].

The form of mushroom cell simulated is essentially that shown in schematic in Fig. 8.9 and having a heater diameter φ (radius r_h), a phase-change (GST) layer thickness of L_c and with the horizontal extent (width) of the phase-change layer being w. Typical outputs from the simulations are shown in Fig. 8.12. Here we see the form of the amorphized and subsequently recrystallized regions as the cell is scaled down in size (specifically as the heater is scaled down from 100 nm in diameter to 6 nm while keeping the scaling factors of L_c/φ and w/L_c constant) and for electrical RESET and SET pulses of 2.5 V/40 ns and 1.5 V/100 ns, respectively (and with rise/fall times of 15/5 ns and 30/30 ns, respectively). It is clear from Fig. 8.12 that both RESET and SET processes were able to be carried out successfully for heater sizes of only 6 nm. In addition, the RESET currents were dramatically reduced as the cell scaled down in size, from over 1 mA for the largest cell (100 nm heater diameter) to less than 20 μA for the smallest, i.e. 6 nm heater diameter, cell. For the range of heater contact diameters shown in Fig. 8.10 (i.e. 56 nm–10 nm), the RESET current values for the physically realistic simulations of Fig. 8.12 varied from approximately 485–40 μA, in good agreement with the simpler (thermal and electrical only) simulations of Fig. 8.10. Note however that to achieve successful amorphization for heater sizes below 15 nm in the results of Fig. 8.12, we had to include a thermal barrier layer (see Sect. 8.4.1) at the top electrode to enable the melting temperature to be reached for reasonable voltages. It should also be pointed out that the resistance window (i.e. R_{RESET}/R_{SET}) decreased as the cell shrunk, and the SET resistance increased to relatively large values, both disadvantageous in terms of the veracity and speed of the readout process.

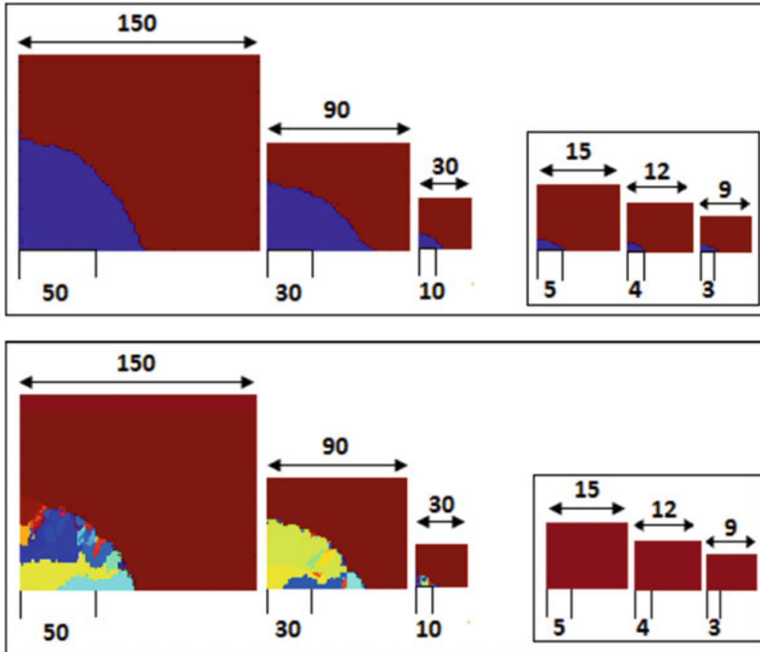


Fig. 8.12 (top) Simulated amorphization of a PCM mushroom-type cell at the end of the RESET process for various levels of device scaling (numbers at the top and bottom of plots show the GST width w_c and the heater radius r_h, respectively, both in nanometres); blue and brown colours show amorphous and crystalline phases, respectively. (bottom) Recrystallization of the amorphized regions shown at the end of the SET process; multiple colours reflect crystalline grains with different orientations (note that for the smallest heater sizes, the cell recrystallizes by growth only from the pre-existing amorphous to crystalline boundary)

The various scaling studies discussed in this section show that PCM mushroom-type cells do indeed have the potential to be scaled down to very small dimensions and that in doing so, high programming currents and other device challenges (such as thermal crosstalk) can be overcome by using both isotropic and non-isotropic scaling approaches. In the following section, we review some of the progress that has been made experimentally in realizing such scaled-down mushroom cells.

8.3.1.2 Experimental Scaling of PCM Mushroom-Type Cells

The first widespread reports on the performance of commercially viable PCM mushroom-type cells appeared for the (F=) 180 nm node in the late 1990s and early 2000s. Not surprisingly for such relatively large cells, key operating characteristics such as the RESET current were relatively poor, for example, a RESET current of around 1 mA in [47]. However, even as early as 2003, experimental mushroom cells with much improved performance were demonstrated, such as a

65 nm node cell with 50 μA RESET current by Pirovano et al. [40]. Interestingly, Pirovano in the same work also showed a factor of ×2 reduction by moving from a mushroom cell design to its effective 'pore' cell complement, showing that switching current and power reductions could be achieved by judicious device engineering as well as via simple scaling (note that we will return to pore cells in Sect. 8.3.2.2).

Product-style PCM memories fabricated at the 90 nm mode were first revealed around 2006, with, as expected, significantly improved performance as compared to 180 nm node devices. For example, Pellizzer et al. [44] demonstrated arrays of mushroom-type PCM cells having RESET currents of around 700 μA for a heater contact area of approximately 3000 nm^2 (i.e. a heater diameter of around 60 nm) and a complete cell area (including the selector devices, etc.) of $12F^2$. More recently IBM have demonstrated 90 nm node mushroom-type PCM cells that can successfully be used to store and recover up to 3 bits (or eight separately differentiable resistance levels) per cell (see Fig. 8.13), significantly enhancing the potential storage capacities of phase-change memories [48]. Interestingly the same 90 nm node mushroom-type cells have also been used to demonstrate some of the other remarkable functionalities offered by phase-change devices, including arithmetic processing [49] and even the development of neuronal mimics for brain-like processing [50].

Beyond (i.e. smaller than) the 90 nm node, the simple mushroom cell has not been extensively utilized for product-oriented PCM memory development; rather variations of it, such as the μTrench [51], dash-type [52, 53] and pore cells, have predominated (see, e.g. [54]). Such cell designs are discussed in the following sections.

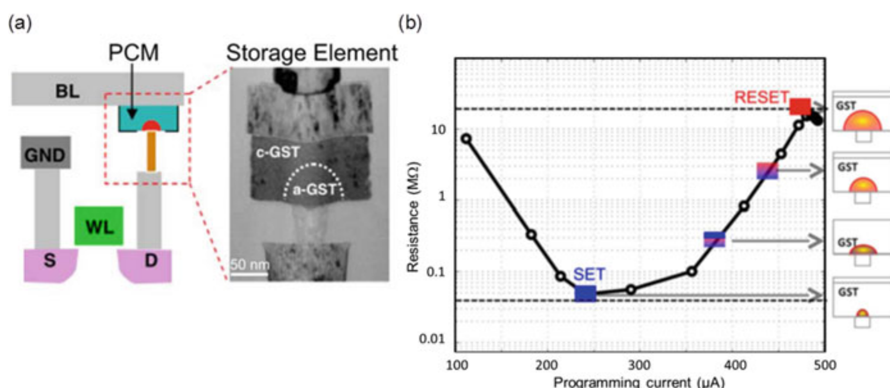


Fig. 8.13 (**a**) A 90 nm node mushroom-type PCM as developed by IBM to deliver multilevel storage capability of up to 3 bits per cell. (**b**) Programming current versus resistance curve for multilevel operation showing use of partial RESET to provide multiple resistance levels (Both reprinted with permission from [48]; copyright IEEE 2016)

8.3.2 Scaling in Other Product-Oriented PCM Cell Designs: The μTrench, Dash-Type, Pore and Crossbar Cells

8.3.2.1 The μTrench and Dash-Type Cell

The μTrench structure reduces the switching volume in the phase-change cell by depositing the phase-change layer into a lithographically defined trench that intersects with a thin, vertical heater electrode [51]. At the 90 nm node, such cells could be programmed (RESET) with a current of around 400 μA, compared to around 700 μA for a mushroom cell fabricated at the same technology node. A later variant of the μTrench, the so-called 'wall' structure, was successfully used to develop a 1Gbit PCM array at the 45 nm node (with an effective cell size of $5.5F^2$) and a RESET current of 200 μA [55].

The dash-type cell (see Fig. 8.14), developed by Samsung, is fabricated by first forming a sidewall electrode, recessing the electrode vertically and filling the recessed volume with phase-change material and then (after removal of any phase-change material from the top surface) depositing the top electrode. In this way the active volume of phase-change material can be dimensionally constrained to very small volumes (since it is confined to the recessed volume), with a resulting decrease in switching currents, as shown in Fig. 8.14a [52]. Indeed, dash cells with phase-change volumes as small as $7.5 \times 22 \times 30$ nm (width $\times$ depth $\times$ height; see Fig. 8.14b) were successfully fabricated at the 20 nm node and showed RESET currents below 100 μA, along with other excellent performance characteristics

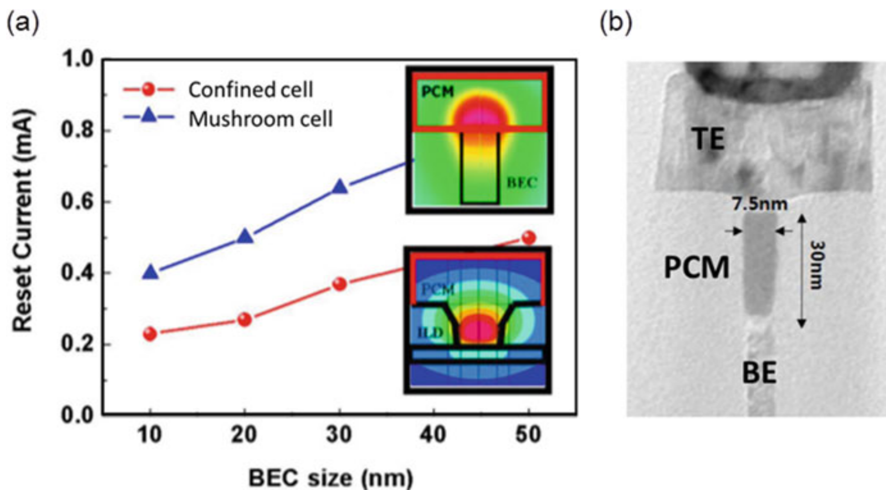


Fig. 8.14 (**a**) Simulated RESET current amplitude as a function of the size of the bottom electrical contact (BEC) for a mushroom-type cell (planar structure) and a confined cell, showing expected reduction in RESET current for the latter (Reprinted with permission from [52]; copyright IEEE 2008). (**b**) TEM cross-sectional image of a dash-type cell, as developed by Samsung (Reprinted with permission from [56]; copyright IEEE 2011)

including around 10^{11} memory cycles without failure and projected lifetimes of more than 10 years at 85 °C [53, 56]. Dash-type PCM arrays with a total capacity of 8 Gbits and a cell size of $4F^2$ at the 20 nm node were also demonstrated.

8.3.2.2 The Pore-Type Cell

The pore-type cell, as shown in Fig. 8.15, is another approach that has been used to confine the volume of phase-change material that undergoes switching. By limiting the phase-change material that is in contact with the bottom heater electrode to a small 'pore' region, the size of the hot spot within the phase-change layer is effectively determined by the pore size (see Fig. 8.15b), with a consequent reduction in active switching volume and in switching current and energy (see Fig. 8.15c).

So far we have concerned ourselves in this chapter with the effect of device scaling on switching current requirements, specifically the RESET (or melting) current, since it is this that determines the energy/power consumption of the device and the current driving requirements of the individual cell-selector electronics (and whose size also depends on the current they are required to supply). However, it is also important to consider other switching characteristics and how these are affected by the scaling of the cell size. In particular, the switching speed is an important criterion that should not be overlooked. Since amorphization processes are *de facto* fast – once we have heated the active region up to the melting temperature, it has to be cooled down at very fast rates, typically >40 Kns^{-1}, in order to freeze into the amorphous phase (rather than recrystallize) – it is the crystallization process that dominates the overall speed of PCM devices. The rate of crystallization is determined by the crystal nucleation rate and the crystal growth rate (both of which are of course functions of temperature), and clearly the smaller the volume of amorphized material that is required to be recrystallized in a cell, the quicker the recrystallization process will be (assuming that said volume contains or

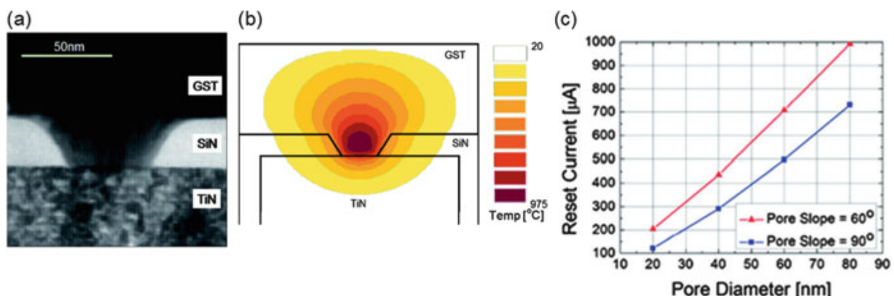


Fig. 8.15 (**a**) TEM cross-sectional image of a GST pore-type cell. (**b**) Simulation of temperature distribution in GST layer during RESET process, showing successful confinement of the heated volume. (**c**) Simulated effect of pore diameter and slope of SiN sidewall on RESET current (All reprinted with permission from [54]; copyright IEEE 2007)

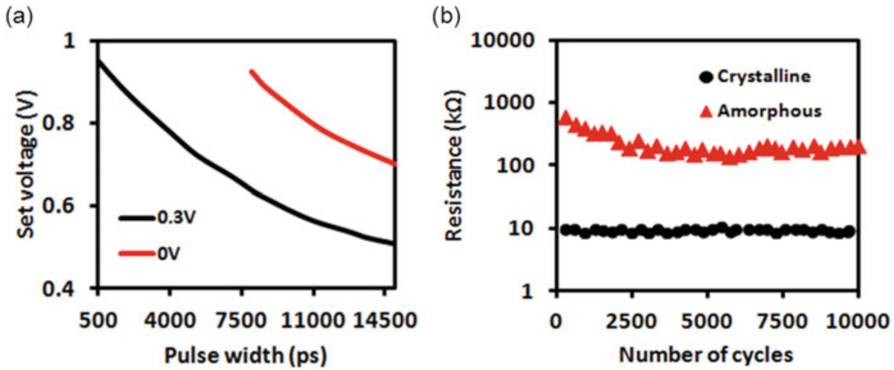


Fig. 8.16 (**a**) SET voltage versus pulse width for pore-type cells of 50 nm in diameter with and without the aid of a 0.3 V 'incubation field'. Without the incubation field, the shortest SET pulse that switched the cell to the low-resistance state was 10 ns, whereas with incubation it was 500 ps. (**b**) Shows successful repeated switching of the cell in (**a**) with 500 ps pulses for both SET and RESET (Both from [58])

can generate sufficient crystal nuclei during the crystallization process and/or borders a region that is already crystallized and can grow into the amorphous volume).

The interplay between active phase-change volume and switching speeds has thus been studied by several researchers using pore-type cells. For example, Wang et al. [57] fabricated GST pore-type cells with bottom contact sizes (diameters) ranging from 500 nm down to 19 nm and investigated the switching speed as a function of cell size. Extremely fast, for PCM, switching was obtained for the smallest cells, with 2.5 ns SET and 400 ps RESET pulses (of 0.8 V and 4.5 V, respectively) reported. Subsequently, by exploitation of an incubation pulse prior to the main switching pulse (essentially a form of preheating of the cell so as to control nucleation) and again using pore-type cells, Loke et al. [58] were able for the first time to demonstrate successfully PCM cell crystallization using sub-ns, specifically 500 ps, SET pulses (see Fig. 8.16). Such an improvement in switching speed with device scaling provides a very welcome added benefit and opens up the possibility of using PCM devices in areas not previously thought practicable, such as in DRAM replacement.

8.3.2.3 Crossbar Cells

Crossbar cells (Fig. 8.17) have gained significant interest in recent years as they provide the possibility for 3D memory layer stacking and high scalability [59–61]. In a crossbar cell, the active material is placed at the intersections of crosswise (orthogonal) patterns of metal lines, the so-called bit line (BL) and word line (WL), and SET/RESET excitations placed on a bit line/word line pair are passed to the phase-change cell via an integrated selector device (ideally a diode to limit area

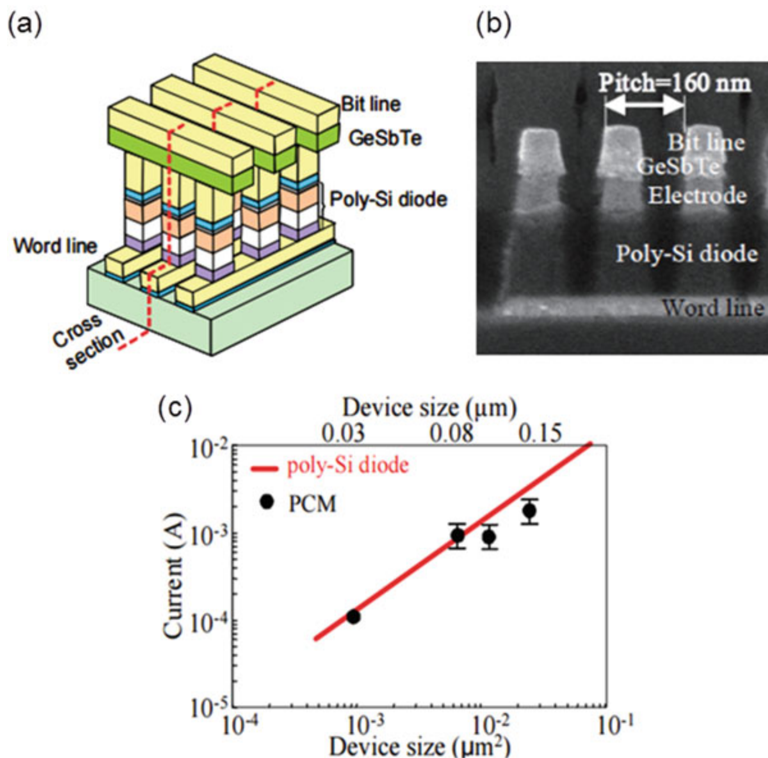


Fig. 8.17 (**a**) Schematic of a crossbar-type PCM structure with integrated Si-diode selector. (**b**) TEM cross-sectional image of crossbar device. (**c**) RESET current scaling characteristics of a PCM crossbar device showing a tenfold reduction in current as the contact size shrinks from 150 nm to 30 nm (All reprinted with permission from [60]; copyright IEEE 2009)

footprint) (see Fig. 8.17a). The active volume of phase-change material is essentially determined by the size of the electrodes, with an approximately tenfold reduction in RESET current (from >1500 μA to 160 μA) having been reported [60] as contact sizes were reduced from 150 nm to 30 nm (see Fig. 8.17c). A stackable and scalable crossbar cell using a chalcogenide threshold switch (also known as ovonic threshold switch or OTS) as the selector, rather than an integrated Si diode, has also been demonstrated, opening up the exciting prospect of an entirely chalcogenide-based crossbar-type memory [59].

8.3.3 Scaling in Research-Oriented PCM Cell Designs

If we move away from the constraints imposed by product-oriented cell designs, there is much more scope to fabricate PCM cells that can truly test experimentally

the possible limits of scalability for phase-change memories. In addition, by moving to a planar, horizontal geometry (which would be unattractive for product-oriented designs due to a large areal footprint), we also open up the possibility for in situ microscopic study of device performance during switching processes, yielding very useful information linking physical and electrical behaviours (see, e.g. [62]). Such planar geometry devices usually go by the name of line, lateral or bridge cells, and scaling studies using such designs are now discussed.

8.3.3.1 Lateral, Line and Bridge Cells

The lateral cell (see Fig. 8.18) is very simple in concept: a thin phase-change layer of thickness H and width W sits between two planar electrodes separated by a distance L, as shown in Fig. 8.18a. The line cell is a slight variant of the lateral cell in so much as the width of the phase-change material lying between the electrodes is reduced lithographically so that the electrodes are connected by a line of phase-change material. The bridge cell can be viewed as a variant of the line cell in so much as the line of phase-change material sits on top, rather than between, the electrodes, as also shown schematically in Fig. 8.18c. The use of such planar cell designs allows us to severely restrict, using high-resolution lithography and ultra-thin phase-change films, the volume of material involved in switching, providing a most useful vehicle for the study of scaling properties.

One of the first lateral PCM cell designs to be demonstrated was by Bolivar et al. who reported RESET currents of around 140 μA [63, 64] for lateral cells having a width and length of around 90 nm and a phase-change layer thickness of 30 nm (and with GST as the active material). Lankhorst reported similar values for RESET currents in line cells fabricated using a doped SbTe phase-change alloy and having a cross-sectional area (W x H) of around 15 nm^2 [65]. Chen et al. [66] by using bridge cells with ultra-thin GeSb phase-change layers down to 3 nm in thickness again demonstrated the clear dependence of contact area on RESET current, achieving sub-100 μA values for the smallest cells.

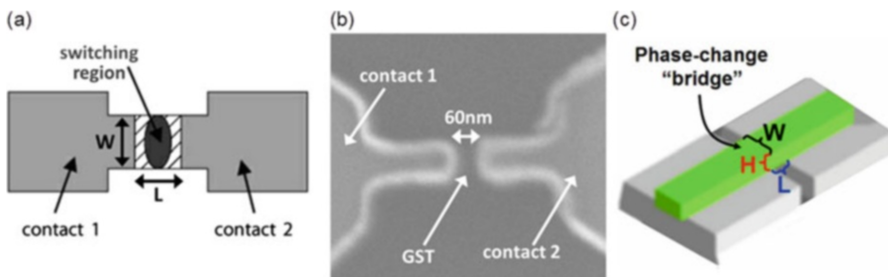


Fig. 8.18 (**a**) Schematic of a PCM lateral cell and (**b**) SEM image of an e-beam lithographed lateral cell having electrodes of width, W, of approximately 90 nm and separated by a distance, L, of around 60 nm. (**c**) Schematic of the basic structure of a PCM bridge cell

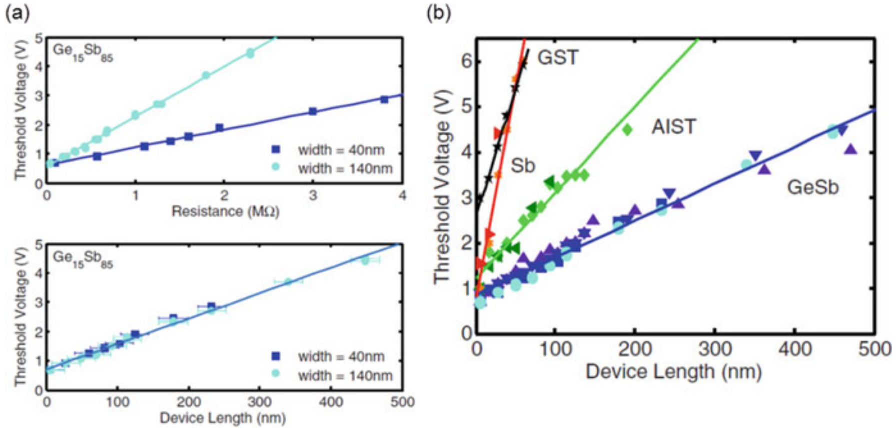


Fig. 8.19 (**a**) The scaling of threshold switching voltage for bridge cells of various sizes and (**b**) the scaling of threshold voltage with bridge cell length for cells made from various phase-change materials (Both reprinted with permission from [67]; copyright AIP Publishing LLC 2009)

Bridge cells of the basic form shown in Fig. 8.18c were also used by Krebs et al. [67] to study the effects of geometric scaling on switching voltage, specifically on the threshold voltage (V_{TH}) at which PCM devices suddenly start to conduct while still effectively in the amorphous phase. It was found that V_{TH} varied linearly with the resistance of the bridge cell and linearly with the length, L, of the cell, as shown in Fig. 8.19. Although resistance also depended on the cell width, W, the threshold voltage did not. Such results can be explained by assuming that V_{TH} is determined by a threshold electric field E_{TH} such that

$$V_{TH} = E_{TH}L = \frac{E_{TH}WH}{\rho}R \tag{8.16}$$

where $R = \rho L/(WH)$ and ρ is the resistivity of the phase-change material. The above equation explains why V_{TH} depends linearly on R and varies with W but, assuming E_{TH} is fixed for a particular phase-change material, why V_{TH} scales with cell length L for any width W. By fabricating devices of various lengths and widths for a range of phase-change materials, including GST, AIST, GeSb and even elemental Sb, Krebs et al. showed that the relation shown in Eq. (8.16) held for all cases and were thus able to determine the threshold electric fields for these materials (which were found to be approximately 56, 19, 8.1 and 94 MVm^{-1}, respectively).

8.3.3.2 ‘Ultimate’ Scaling in CNT-Based Cells

The lateral, line and bridge-type cells discussed in the previous section utilized high-resolution e-beam lithographic patterning and allowed the examination of PCM scaling down to contact areas of a few tens of square nanometres using

'conventional' metal electrodes. To reduce the contact area to the sub-10 nm^2 region, and to probe the 'ultimate' scaling capabilities of phase-change memories, several authors have reported the use of carbon nanotube (CNT) electrodes having diameters typically in the range of 1–6 nm [68–70]. For example, in seminal work by Xiong et al. [68], working phase-change memories were fabricated by forming (via electrical breakdown) a nanogap along the length of a carbon nanotube and then filling the nanogap with a thin phase-change layer (10 nm of GST in this case; see Fig. 8.20). Although the phase-change layer was not laterally confined, i.e. it extended laterally far beyond the diameter of the CNT itself, the written bits themselves appeared to be well confined (due to thermal effects) to lateral sizes similar to the CNT diameter (see Fig. 8.20a), meaning that the effective bit volumes were as small as a few hundred cubic nanometres.

By using nanotubes of various diameters, Xiong et al. were able to determine the scaling properties of their CNT-based PCM devices, finding RESET currents as low as 5 μA (and SET currents below 1 μA) for the smallest devices having a contact diameter of around 2–3 nm (see Fig. 8.20b). In subsequent work, Xiong et al. [70] developed a fabrication technique (involving a form of phase-change nanowire being deposited into a nanotrench formed around the CNT) to limit the dimensions of the phase-change layer in the lateral direction perpendicular to the CNT nanogap. This reduced the effective contact area still further (to ~2.5 nm^2, diameter ~ 1.8 nm) as well as any lateral spread of heat (as compared to [68]), resulting in RESET currents as low as 1.6 μA being achieved. Liang et al. [69] were also able to obtain ultralow programming currents in CNT-based cells having contact areas of ~2.5 nm ^{2} but using a crossbar-type device configuration more closely related to product-like device architectures (see Fig. 8.21).

Along with the dramatic reductions in programming currents achieved using CNT-based PCM cells also come vast reductions in programming power and energy. For example, the nanogap devices reported by Xiong et al. [68] achieved

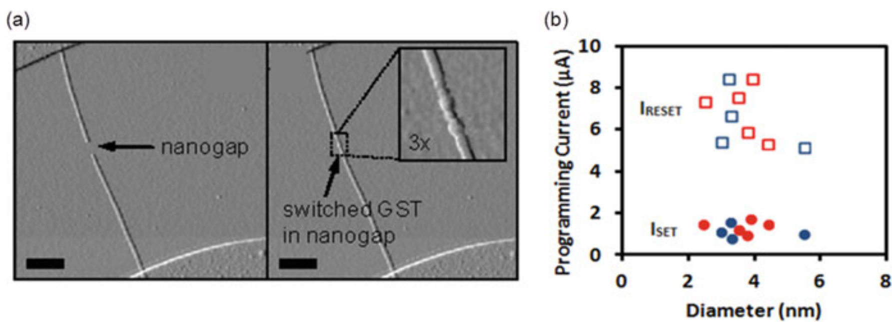


Fig. 8.20 (**a**) AFM (atomic force microscope) images of a CNT-based nanogap PCM cell before (left) and after (right) filling with GST (the scale bars are 500 nm) (Reprinted with permission from [68]; copyright The American Association for the Advancement of Science 2011). (**b**) Ultralow programming currents achieved for several CNT-based nanogap PCM devices having ultrasmall contact diameters (blue and red dots correspond to devices where the formation of the nanogap was carried out in air and Ar, respectively) (Adapted from [68])

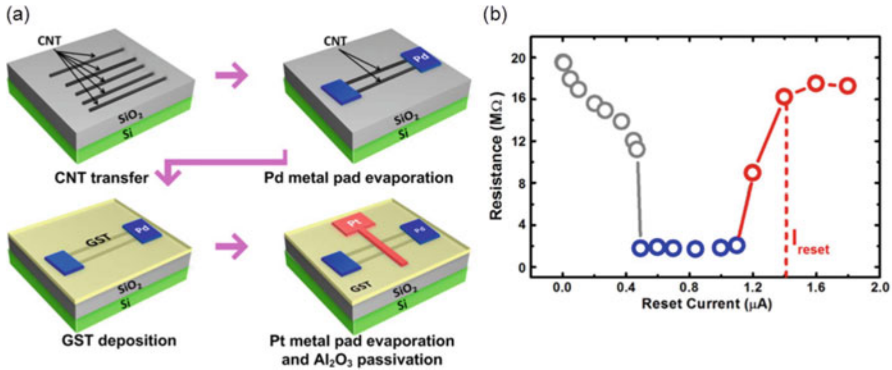


Fig. 8.21 (**a**) Fabrication process for CNT-based crossbar-type PCM cells. (**b**) The programming R-I curve for a device of the type shown in (**a**) and having an effective contact diameter of approximately 1.8 nm; ultralow RESET currents of ~1.4 μA are achieved (and SET currents of <0.5 μA).] (Both reprinted with permission from [69]; copyright IEEE 2012)

RESET powers of a few μW (cf. ~ mW programming power for product-like devices) and RESET energies of around 100 fJ for 20 ns RESET pulse lengths. Further reductions in switching energies would be expected for smaller bits and faster pulses. Indeed, Xiong et al. speculated that if the nanogaps were to be reduced in length to around 5 nm, then the active switching volume in their CNT-based cells would be reduced to around 20 nm^3, and the resulting SET and RESET energies would fall as low as 5×10^{-18} and 2×10^{-17} J, respectively (i.e. 5 aJ and 20 aJ) – truly miniscule values. And, even if such small switching volumes (~ 20 nm^3) were to be achieved in some 'ultimately scaled' device, these would still be substantially larger than the 2 nm diameter (~4 nm^3) nanoparticles that have been shown experimentally to successfully crystallize (see, e.g. Fig. 8.6b and [33]) and would still be substantially larger than the theoretical minimum crystallite size predicted by classical nucleation theory (e.g. ~ 0.5–1 nm^3 for GST; see Fig. 8.7a and [37]) and by atomistic simulations (e.g. 1.5–3 nm^3 for GST; see Fig. 8.7b and [38]), meaning that even smaller working cells (at least of the research-oriented type) might be imagined.

8.4 The Role of Material Innovation and Material Engineering in Scaling

8.4.1 Thermal Engineering and Scaling

In the previous sections, we have considered the effects of PCM cell size and cell design on the device operating characteristics as a function of scaling. A key benefit of scaling has been seen to be the reduction of RESET current (and hence device operating power/energy) as cell sizes shrink. However, since the phase

transformation process in PCM cells is essentially thermally driven, by judicious thermal engineering of the cell, we might also expect to realize benefits in terms of a reduction in programming current. Essentially we would like to thermally engineer the cell so as to retain, in the active switching volume, as much heat as possible (maximizing the temperature rise per unit excitation current) while at the same time ensuring that the melted phase can cool fast enough (i.e. greater than several tens of Kelvin per nanosecond) for resolidification to occur into the amorphous phase. These two thermal requirements are in essence in opposition to each other, but nonetheless, by careful design, we can indeed fulfil both requirements and yield certain device performance improvements.

Since most heat will be lost from a PCM cell via the usually high thermal conductivity electrodes, one approach to thermally engineer a reduction in programming current is to modify in some way such electrodes to reduce their thermal conductivity while at the same time maintaining adequate electrical contact. An approach that achieves this has already been used in the simulations shown in Fig. 8.12; here, to achieve amorphization within the active region for reasonable applied voltages, we used, for the smallest cells (as shown in the inset of Fig. 8.12a), a multilayered TiN/W top electrode (rather than a single-layered TiN or W electrode) for which the thermal conductivity can be reduced to below 1 Wm^{-1} K^{-1} (see [71]) by varying the number and thicknesses of the alternating TiN and W layers. The electrical resistivity of multilayered TiN/W electrodes was, however, not so different to that of single-layer TiN electrodes, though significantly higher than that of W alone (e.g. a 50-nm-thick electrode of alternating 5 nm TiN/W layers had a resistivity of 1.1×10^{-4} Ω.cm as compared to 1.0×10^{-4} Ω.cm and 0.6×10^{-5} Ω.cm for single-layer TiN and W electrodes, respectively – see [72]).

Another thermal engineering approach is to modify the thermal environment and/or thermal conductivity of the bottom heater contact. For example, Wu et al. [73] reported that thermal confinement of the TiN bottom electrode by surrounding it with a much lower thermal conductivity TaN region could result in a $\times 10$ reduction in RESET current for the same cell size. Graphene has also proved to be a useful thermal barrier layer in PCM cells, with graphene-PCM devices being shown to have ~40% lower RESET current as compared to notionally identical control devices without the graphene layer (see Fig. 8.22 [74]). The thermal effects of graphene in this instance were attributed to a very considerable increase in thermal boundary resistance caused by its insertion (equivalent in fact to the thermal resistance of around 10–15 nm of GST, simply from an atomically thin graphene layer).

8.4.2 Phase-Change Material Engineering and Scaling

Another approach to improving the operating characteristics of PCM devices is, of course, to use different/improved phase-change materials. In most of this chapter,

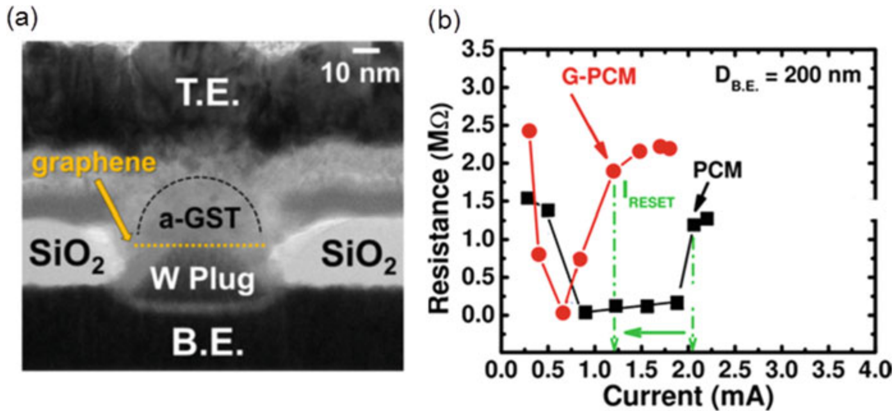


Fig. 8.22 (**a**) TEM cross-sectional image of a PCM cell with a graphene thermal barrier layer placed between the bottom electrode and the active region of the phase-change layer and (**b**) the resulting programming R-I curves for the graphene-modified cell (G-PCM) and a control cell with no graphene layer (PCM) clearly showing the reduction in RESET current resulting from use of the graphene layer (Both reprinted with permission from [74]; copyright American Chemical Society 2015)

we have concentrated mainly on the use of GST ($Ge_2Sb_2Te_5$) so as to provide a 'level playing field' when comparing the performance of devices subject to different amounts of scaling. We are however at liberty to use a wide variety of phase-change materials in the place of GST, such as N-doped GST, C-doped GST, $Ge_2Sb_1Te_2$, GeTe, AIST, Sb_2Te, GeSb, etc. (the list is almost endless; see Fig. 8.2a to get an idea of the range of even just telluride-based phase-change alloys). Each of these compositions usually has some trade-off between various key material properties (e.g. crystallization temperature, melting temperature, crystallization speed, thermal conductivity, electrical conductivity, etc.), and this translates into a trade-off between key device operating characteristics (e.g. crystallization (SET) speed, temperature stability, programming current, endurance, retention) such that, in terms of alloys, GST and its close variants continue to be the materials of choice of PCM applications (particularly the so-called golden composition of Ge-doped $Ge_2Sb_1Te_2$; see Fig. 8.2b and, e.g. [4]). However, as discussed briefly in Sect. 8.2.2, there does exist the intriguing possibility in some instances of inducing phase-switching in chalcogenide materials without the need for energy-consuming melting processes. While in Sect. 8.2.2 non-melt-quenched amorphization was discussed in terms of defect engineering in nanowires [25], a possibly more practicable approach to achieving non-melt-quenched phase-switching in a device configuration is via the use of the so-called phase-change superlattice-like (SLL) materials, as outlined below.

Phase-change SLL structures typically comprise multiple atomically thin layers of alternating phase-change compositions, e.g. GeTe and Sb_2Te_3. They first began to appear in the early 2000s (see, e.g. [75]) as an attempt to combine in a single film the attractive properties of different phase-change alloys (e.g. GeTe has relatively high crystallization and melting temperatures leading to good high-temperature stability, but it has a relatively slow crystallization speed; Sb_2Te_3, on the other hand, has relatively low crystallization and melting temperatures leading to poor thermal stability but has a high crystallization speed). Indeed, PCM cells with high switching speed (1 ns SET pulses) and low programming currents as compared to GST equivalents were reported using such an approach [76].

More exciting revelations came however with the work of Simpson et al., where a non-melting transition between two solid phases in GeTe/Sb_2Te_3 SLL devices was observed [77]. Such a non-melting phase-transition is expected to lead to reduced power consumption (since the phase-change layer does not have to be heated to the melting temperature and the lack of a bulk amorphization process reduces entropic switching energy losses) and improved endurance, i.e. the number of SET/RESET cycles before failure (due to reduced thermal and mechanical stresses of lower switching temperatures). Indeed, Simpson et al. reported RESET currents reduced by around 40% for SLL devices as compared to GST equivalents, along with significant improvements in endurance [77] (see Fig. 8.23). While the precise mechanism for the phase-switching process in SLL films is still a matter for some debate [78–80], and while the fabrication of atomically thin SLL films is considerably more demanding than that of

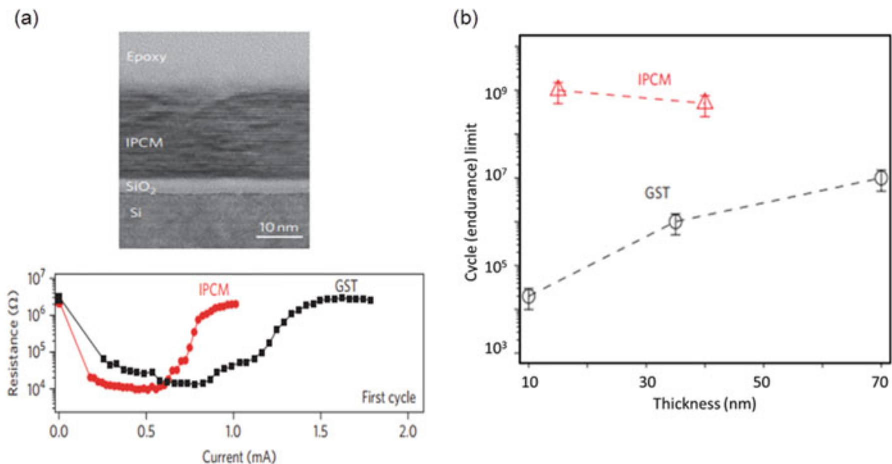


Fig. 8.23 (**a**) HR-TEM image of a typical as-grown $(GeTe)_2$ $(Sb_2Te_3)_4$ SLL film (top) and R-I curves for SLL (IPCM) cell as compared to an identical cell fabricated using GST-225 alloy (bottom). (**b**) Maximum number of SET-RESET switching cycles (i.e. the endurance) for IPCM and GST-alloy cells, plotted as a function of material thickness (All reprinted with permission from [77]; copyright Nature Publishing Group 2011)

conventional alloy films, the prospect of non-melting phase-switching for future PCM devices is an attractive one.

8.5 Summary and Outlook

It is clear from our discussions in the previous sections that we have gone a long way towards answering many of our questions in terms of how far PCM devices might be scaled down in size and still function adequately. Indeed as PCM cells become smaller, they function more than adequately; in fact they function very significantly better than larger cells in terms of device operating parameters, in particular programming currents, powers and energies all reduce, and switching speed is increased. The scaling down of PCM cell size is notably beneficial with regard to the reduction of the RESET current, which is probably *the* critical device characteristic in terms of commercial viability. In fact we have seen in Sect. 8.3 that RESET current depends strongly on the electrode contact size and is predicted to vary as the inverse of contact diameter for isotropic scaling (see Sect. 8.3.1.1) and by more than this for non-isotropic scaling (see, e.g. Fig. 8.10b). We have also seen, not unsurprisingly, that the device design has a considerable role to play in minimizing RESET current at any particular technology node (compare, e.g. the confined versus planar cell geometries shown in Fig. 8.14a). Material engineering can also play a significant role in reducing the programming currents, in particular via the control of thermal boundary resistances (and hence overall thermal resistances) at interfaces, as exemplified via the use of a graphene thermal barrier layer (as in Fig. 8.22). The type of phase-change material used itself will also play a significant role, since electrical, thermal and phase-change properties vary from composition to composition. Indeed, the exploitation of a non-melting resistive switch via the use of novel phase-change superlattice-like materials, or by defect engineering, can also reduce the RESET current requirements.

The improvements in RESET current achieved via all the above approaches are summarized in Fig. 8.24. The beneficial effects of scaling are plain to see: the largest cells show RESET currents of around 1 mA, reducing to 50–100 μA for product-like PCM cells having contact sizes of around 20 nm, right down to just over 1 μA for the smallest cells fabricated using CNT electrodes and having contact sizes of only 2 nm or so. Of course RESET current, although a primary concern in terms of PCM development, is not the only device characteristic of interest in terms of phase-change memory operation. We need also to consider, in particular, switching power/energy, switching speed, retention time and endurance. Switching power is also a major beneficiary of scaling, since, as we saw in Sect. 8.3.1.1, the power consumed during RESET scales in the same way as the RESET current, at least for isotropic scaling (see Eq. 8.15). The switching energy will depend on the duration of the RESET pulse, but since by necessity this needs to be short (to ensure melt quenching into the amorphous phase) and RESET has been demonstrated with sub-ns pulses already, we can expect ultralow RESET pulse energies in the femto-

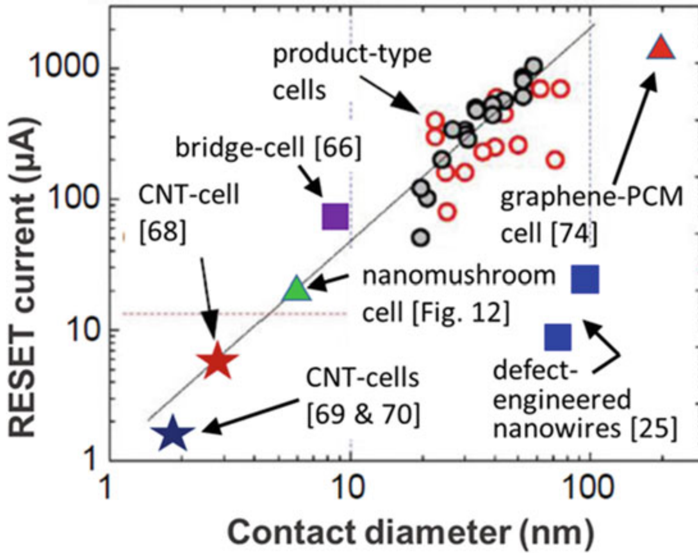


Fig. 8.24 The variation of RESET current as a function of contact diameter for various product-oriented and research-oriented PCM cells

Joule region for the smallest of cells (see Sect. 8.3.3.2). Turning to switching speed, this, as we have also discussed in Sect. 8.3.2.2, is limited by how fast the amorphized volume can be recrystallized (at SET process temperatures), and the smaller the volume, the quicker such recrystallization will be completed (see, e.g. Fig. 8.16a). Retention on the other hand is likely to be adversely affected by size scaling, since retention is also essentially determined by how long an amorphized cell takes to recrystallize at ambient temperatures. In terms of endurance, even the truly nanoscale lab-fabricated research-oriented CNT-based PCM cells have shown a fairly good number of switching cycles (e.g. ~ 1000 in [70]), giving us confidence that properly manufactured nanosized cells could yield very acceptable endurance values. And, if a way can be found to exploit in a commercially viable way in nanosized cells a switching process based on non-melt-quenched amorphization/phase transitions, then very significant improvements in endurance would be expected.

Finally, we summarize in Table 8.2 PCM device performance attributes for the main product-oriented and research-oriented cell designs. It is clear from this, and from Fig. 8.24, that if we can successfully manufacture PCM devices down to single-nanometre technology nodes, and do so in a commercially viable way, then PCM device technology is likely to have a bright future.

Table 8.2 A comparison of performance attributes for various product-oriented and research-oriented PCM cells fabricated at various scaling points

PCM cell type	Technology node F (nm) and total cell area (F^2, inc. selector)	Contact dia. (nm) and area (nm^2)	RESET current (μA)	SET speed (ns)	Endurance (# cycles)	Switching power (μW) RESET)	Notes
Mushroom	180 (Na)	~ 90(~ 6000)	> 1000	50	~ 10^{12}	~ 7000	See [47]
Mushroom	90 ($12F^2$)	~ 60 (~3000)	~ 700	100	> 10^8	~ 2200	See [44]
Nano-mushroom	>5 (Na)	5 (~ 20)	~ 15	< 50 ns	NA	~ 1	From results of Fig. 8.12
μTrench	180 ($10F^2$)	NA	600	100	> 10^{11}	~ 3600	See [51]
μTrench	90 ($12F^2$)	~ 20 (~400)	400	100	> 10^8	~ 600	See [44]
μTrench (wall)	45 ($5.5F^2$)	NA	200	NA	>10^8	~ 400	See [55]
Dash-type	~20 (<$4F^2$)	~7.5 × 22 (127–165)	80–100	50	~ 10^{11}	NA	See [52, 53]
Pore type	180 (Na)	~ 40 1250	250	80	>10^5	~ 1500	See [54]
Pore type	Na (Na)	~50 (~2000)	NA	0.5, 10	> 10^4	NA	See [58] (0.5 ns with incubation, 10 ns without)
Crossbar	80 ($4F^2$)	30 (~900)	160	1000	NA	NA	See [60]
CNT type	~2 (Na)	<2 (~ 2.5)	~ 1.5	300	10^2–10^3	0.08–0.2	See [68–70]

References

1. G. Moore, Cramming more and more components onto integrated circuits. Electronics **38**, 8 (1965)
2. A.L. Lacaita, A. Redaelli, The race of phase change memories to nanoscale storage and applications. Microelectron. Eng. **109**, 351–356 (2013)
3. G.W. Burr, M.J. Breitwisch, M. Franceschini, D. Garetto, K. Gopalakrishnan, B. Jackson, B. Kurdi, C. Lam, L.A. Lastras, A. Padilla, B. Rajendran, Phase change memory technology. J. Vac. Sci. Technol. B **28**(2), 223–262 (2010)
4. G.W. Burr, M.J. Brightsky, A. Sebastian, H.-Y. Cheng, J.-Y. Wu, S. Kim, N.E. Sosa, N. Papandreou, H.-S. Lung, H. Pozidis, E. Eleftheriou, C.H. Lam, Recent progress in phase-change memory technology. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **6**(2), 146. [Online]. Available:http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7453199 (2016). https://doi.org/10.1109/JETCAS.2016.2547718
5. G. Bruns, P. Merkelbach, C. Schlockermann, M. Salinga, M. Wuttig, T.D. Happ, J.B. Philipp, M. Kund, Nanosecond switching in GeTe phase change memory cells. Appl. Phys. Lett. **95**(4), 043108 (2009)
6. G. Navarro, V. Sousa, P. Noe, N. Castellani, M. Coue, J. Kluge, A. Kiouseloglou, C. Sabbione, A. Persico, A. Roule, O. Cueto.N-doping impact in optimized Ge-rich materials based phase-change memory. In *2016 I.E. 8th International Memory Workshop (IMW)* (pp. 1–4). IEEE 2016
7. S. Souiki, Q. Hubert, G. Navarro, A. Persico, C. Jahan, E. Henaff, V. Delaye, D. Blachier, V. Sousa, L. Perniola, E. Vianello. Reliability study of carbon-doped GST stack robust against Pb-free soldering reflow. In *Reliability Physics Symposium (IRPS), 2013 I.E. International* (pp. MY–8). IEEE 2013
8. H.Y. Cheng, T.H. Hsu, S. Raoux, J.Y. Wu, P.Y. Du, M. Breitwisch, Y. Zhu, E.K. Lai, E. Joseph, S. Mittal, R. Cheek. A high performance phase change memory with fast switching speed and high temperature retention by engineering the Ge x Sb y Te z phase change material. In *Electron Devices Meeting (IEDM), 2011 I.E. International* (pp. 3–4). IEEE 2011
9. D. Lencer, M. Salinga, B. Grabowski, T. Hickel, J. Neugebauer, M. Wuttig, A map for phase-change materials. Nat. Mater. **7**(12), 972–977 (2008)
10. S. Raoux, J.L. Jordan-Sweet, A.J. Kellock, Crystallization properties of ultrathin phase change films. J. Appl. Phys. **103**(11), 114310–114310 (2008)
11. R.E. Simpson, M. Krbal, P. Fons, A.V. Kolobov, J. Tominaga, T. Uruga, H. Tanida, Toward the ultimate limit of phase change in Ge2Sb2Te5. Nano Lett. **10**(2), 414–419 (2009)
12. W. Xiaoqian, L. Shi, T. Chong, Z. Rong, H. Koon, Thickness dependent Nano-crystallization in Ge2Sb2Te5 films and its effect on devices. Jpn. J. Appl. Phys. **46**(4B), 2211 (2007)
13. G.F. Zhou, Materials aspects in phase change optical recording. Mater. Sci. Eng. A **304**, 73–80 (2001)
14. H.C.F. Martens, R. Vlutters, J.C. Prangsma, Thickness dependent crystallization speed in thin phase change layers used for optical recording. J. Appl. Phys. **95**(8), 3977–3983 (2004)
15. S. Raoux, R.M. Shelby, J. Jordan-Sweet, B. Munoz, M. Salinga, Y.C. Chen, Y.H. Shih, E.K. Lai, M.H. Lee, Phase change materials and their application to random access memory technology. Microelectron. Eng. **85**(12), 2330–2333 (2008)
16. Y.C. Her, Y.S. Hsu, Thickness dependence of crystallization and melting kinetics of eutectic Sb 70 Te 30 phase change recording film. J. Non-Cryst. Solids **354**(27), 3129–3134 (2008)
17. J.P. Reifenberg, M.A. Panzer, S. Kim, A.M. Gibby, Y. Zhang, S. Wong, H.P. Wong, E. Pop, K.E. Goodson, Thickness and stoichiometry dependence of the thermal conductivity of GeSbTe films. Appl. Phys. Lett. **91**(11), 111904–111904 (2007)
18. J.L. Battaglia, A. Kusiak, V. Schick, A. Cappella, C. Wiemer, M. Longo, E. Varesi, Thermal characterization of the SiO2-Ge2Sb2Te5 interface from room temperature up to 400 C. J. Appl. Phys. **107**, 044314 (2010)

19. J. Lee, J.P. Reifenberg, Bozorg-Grayeli, L. Hom, Z. Li, S. Kim, M. Asheghi, H.S.P. Wong, K.E. Goodson, Decoupled thermal resistances of phase change material and their impact on PCM devices. In *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010 12th IEEE Intersociety Conference on* (pp. 1–6). IEEE 2010
20. X. Sun, B. Yu, G. Ng, M. Meyyappan, One-dimensional phase-change nanostructure: Germanium telluride nanowire. J. Phys. Chem. C **111**(6), 2421–2425 (2007)
21. S.H. Lee, Y. Jung, R. Agarwal, Highly scalable non-volatile and ultra-low-power phase-change nanowire memory. Nat. Nanotechnol. **2**(10), 626–630 (2007)
22. D. Yu, S. Brittman, J.S. Lee, A.L. Falk, H. Park, Minimum voltage for threshold switching in nanoscale phase-change memory. Nano Lett. **8**(10), 3429–3433 (2008)
23. X. Sun, B. Yu, G. Ng, T.D. Nguyen, M. Meyyappan, III-VI compound semiconductor indium selenide (In2Se3) nanowires: Synthesis and characterization. Appl. Phys. Lett. **89**(23), 233121 (2006)
24. S.W. Nam, H.S. Chung, Y.C. Lo, L. Qi, J. Li, Y. Lu, A.C. Johnson, Y. Jung, P. Nukala, R. Agarwal, Electrical wind force–driven and dislocation-Templated Amorphization in phase-change nanowires. Science **336**(6088), 1561–1566 (2012)
25. P. Nukala, C.C. Lin, R. Composto, R. Agarwal, Ultralow-power switching via defect engineering in germanium telluride phase-change memory devices. Nat. Commun. **7**, 10482 (2016)
26. C.E. Giusca, V. Stolojan, J. Sloan, F. Börrnert, H. Shiozawa, K. Sader, M.H. Rümmeli, B. Büchner, S.R.P. Silva, Confined crystals of the smallest phase-change material. Nano Lett. **13**(9), 4020–4027 (2013)
27. H.R. Yoon, W. Jo, E.H. Lee, J.H. Lee, M. Kim, K.Y. Lee, Y. Khang, Generation of phase-change Ge–Sb–Te nanoparticles by pulsed laser ablation. J. Non-Cryst. Solids **351**(43), 3430–3434 (2005)
28. H.R. Yoon, W. Jo, E. Cho, S. Yoon, M. Kim, Microstructure and optical properties of phase-change Ge–Sb–Te nanoparticles grown by pulsed-laser ablation. J. Non-Cryst. Solids **352**(36), 3757–3761 (2006)
29. G.S. Park, J.H. Kwon, M. Kim, H.R. Yoon, W. Jo, T.K. Kim, J.M. Zuo, Y. Khang, Crystalline and amorphous structures of Ge–Sb–Te nanoparticles. J. Appl. Phys. **102**(1), 013524 (2007)
30. H.S. Choi, K.S. Seol, K. Takeuchi, J. Fujita, Y. Ohki, Synthesis of size- and structure-controlled Ge2Sb2Te5 nanoparticles. *Jpn. J. Appl. Phys.* ***44***(10R), 7720 (2005)
31. S. Raoux, C.T. Rettner, J.L. Jordan-Sweet, A.J. Kellock, T. Topuria, P.M. Rice, D.C. Miller, Direct observation of amorphous to crystalline phase transitions in nanoparticle arrays of phase change materials. J. Appl. Phys. **102**(9), 94305–94305 (2007)
32. Y. Zhang, S. Raoux, D. Krebs, L.E. Krupp, T. Topuria, M.A. Caldwell, D.J. Milliron, A. Kellock, P.M. Rice, J.L. Jordan-Sweet, H.S.P. Wong, Phase change nanodots patterning using a self-assembled polymer lithography and crystallization analysis. J. Appl. Phys. **104**(7), 074312 (2008)
33. M.A. Caldwell, S. Raoux, R.Y. Wang, H.S.P. Wong, D.J. Milliron, Synthesis and size-dependent crystallization of colloidal germanium telluride nanoparticles. J. Mater. Chem. **20** (7), 1285–1291 (2010)
34. G.E. Ghezzi, R. Morel, A. Brenac, N. Boudet, M. Audier, F. Fillot, S. Maitrejean, F. Hippert, Crystallization of Ge2Sb2Te5 nanometric phase change material clusters made by gas-phase condensation. Appl. Phys. Lett. **101**(23), 233113 (2012)
35. G.E. Ghezzi, Material studies for advanced phase change memories: doping, size reduction and interface effect, Doctoral dissertation, Grenoble INP (2013)
36. S. Senkader, C.D. Wright, Models for phase-change of Ge2Sb2Te5 in optical and electrical memory devices. J. Appl. Phys. **95**(2), 504–511 (2004)
37. C.D. Wright, M. Armand, M.M. Aziz, Terabit-per-square-inch data storage using phase-change media and scanning electrical nanoprobes. IEEE Trans. Nanotechnol. **5**(1), 50–61 (2006)
38. T.H. Lee, S.R. Elliott, Ab initio computer simulation of the early stages of crystallization: application to Ge 2 Sb 2 Te 5 phase-change materials. *Phys. Rev. Lett.* ***107***(14), 145702 (2011)

39. A. Sebastian, N. Papandreou, A. Pantazi, H. Pozidis, E. Eleftheriou, Non-resistance-based cell-state metric for phase-change memory. J. Appl. Phys. **110**(8), 084505 (2011)
40. A. Pirovano, A.L. Lacaita, A. Benvenuti, F. Pellizzer, S. Hudgens, Bez, R.. Scaling analysis of phase-change memory technology. In *Electron Devices Meeting, 2003. IEDM'03 Technical Digest. IEEE International* (pp. 29–6). IEEE 2003
41. U. Russo, D. Ielmini, A. Redaelli, A.L. Lacaita, Modeling of programming and read performance in phase-change memories—Part I: Cell optimization and scaling. IEEE Transactions on Electron Devices **55**(2), 506–514 (2008)
42. G.T. Jeong, Y.N. Hwang, S.H. Lee, S.Y. Lee, K.C. Ryoo, J.H. Park, Y.J. Song, S.J. Ahn, C.W. Jeong, Y.T. Kim, H. Horii. Process technologies for the integration of high density phase change RAM. In *2005 International Conference on Integrated Circuit Design and Technology, 2005. ICICDT 2005.* (pp. 19–22). IEEE 2005.
43. U. Russo, D. Ielmini, A. Redaelli, A.L. Lacaita, Modeling of programming and read performance in phase-change memories—Part II: Program disturb and mixed-scaling approach. IEEE Transactions on Electron Devices **55**(2), 515–522 (2008)
44. Pellizzer, F., Benvenuti, A., Gleixner, B., Kim, Y., Johnson, B., Magistretti, M., Marangon, T., Pirovano, A., Bez, R., Atwood, G., 2006. A 90nm phase change memory technology for stand-alone non-volatile memory applications. In *2006 Symposium on VLSI Technology, 2006. Digest of Technical Papers.* (pp. 122–123). IEEE
45. H. Hayat, K. Kohary, C.D. Wright, Can conventional phase-change memory devices be scaled down to single-nanometre dimensions? Nanotechnology **28**, 035202 (2016)
46. P. Ashwin, B.S.V. Patnaik, C.D. Wright, Fast simulation of phase-change processes in chalcogenide alloys using a Gillespie-type cellular automata approach. J. Appl. Phys. **104** (8), 084901 (2008)
47. S. Lai. Current status of the phase change memory and its future. In *Electron Devices Meeting, 2003. IEDM'03 Technical Digest. IEEE International* (pp. 10–1). IEEE 2003
48. A. Athmanathan, M. Stanisavljevic, N. Papandreou, H. Pozidis, E. Eleftheriou, Multilevel-cell phase-change memory: A viable technology. IEEE Journal on Emerging and Selected Topics in Circuits and Systems **6**(1), 87–100 (2016)
49. P. Hosseini, A. Sebastian, N. Papandreou, C.D. Wright, H. Bhaskaran, Accumulation-based computing using phase-change memories with FET access devices. IEEE Electron Device Letters **36**(9), 975–977 (2015)
50. T. Tuma, A. Pantazi, M. Le Gallo, A. Sebastian, E. Eleftheriou, Stochastic phase-change neurons. Nat. Nanotechnol. **11**, 693–699 (2016)
51. F. Pellizzer, A. Pirovano, F. Ottogalli, M. Magistretti, M. Scaravaggi, P. Zuliani, M. Tosi, A. Benvenuti, P. Besana, S. Cadeo, T. Marangon . Novel µtrench phase-change memory cell for embedded and stand-alone non-volatile memory applications. In *VLSI Technology, 2004. Digest of Technical Papers. 2004 Symposium on* (pp. 18–19). IEEE 2004
52. D.H. Im, J.I. Lee, S.L. Cho, H.G. An, D.H. Kim, I.S. Kim, H. Park, D.H. Ahn, H. Horii, S.O. Park, U.I. Chung. A unified 7.5 nm dash-type confined cell for high performance PRAM device. In *2008 I.E. International Electron Devices Meeting* (pp. 1–4). IEEE 2008
53. I.S. Kim, S.L. Cho, D.H. Im, E.H. Cho, D.H. Kim, G.H. Oh, D.H Ahn, S.O. Park, S.W. Nam, Moon, J.T., Chung, C.H.. High performance PRAM cell scalable to sub-20nm technology with below 4F2 cell size, extendable to DRAM applications. In *2010 Symposium on VLSI Technology* (pp. 203–204). IEEE 2010
54. M. Breitwisch, T. Nirschl, C.F. Chen, Y. Zhu, M.H. Lee, M. Lamorey, G.W. Burr, E. Joseph, A. Schrott, J.B. Philipp, R. Cheek. Novel lithography-independent pore phase change memory. In *2007 I.E. Symposium on VLSI Technology* (pp. 100–101). IEEE 2007
55. G. Servalli. A 45nm generation phase change memory technology. In *2009 I.E. International Electron Devices Meeting (IEDM)* 2009
56. M.J. Kang, T.J. Park, Y.W. Kwon, D.H. Ahn, Y.S. Kang, H. Jeong, S.J. Ahn, Y.J Song, B.C. Kim, S.W Nam, H.K. Kang. PRAM Cell Technology and Characterization in 20nm Node Size. In *2011 International Electron Devices Meeting* 2011

57. W.J. Wang, L.P. Shi, R. Zhao, K.G. Lim, H.K. Lee, T.C. Chong, Y.H. Wu, Fast phase transitions induced by picosecond electrical pulses on phase change memory cells. Appl. Phys. Lett. **93**(4), 043121 (2008)
58. D. Loke, T.H. Lee, W.J. Wang, L.P. Shi, R. Zhao, Y.C. Yeo, T.C. Chong, S.R. Elliott, Breaking the speed limits of phase-change memory. Science **336**(6088), 1566–1569 (2012)
59. D. Kau, S. Tang, I.V. Karpov, R. Dodge, B. Klehn, J.A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee. A stackable cross point phase change memory. In *2009 I.E. International Electron Devices Meeting (IEDM)* (pp. 1–4). IEEE 2009
60. Y. Sasago, M. Kinoshita, T. Morikawa, K. Kurotsuchi, S. Hanzawa, T. Mine, A. Shima, Y. Fujisaki, H. Kume, H. Moriya, N. Takaura. Cross-point phase change memory with 4F 2 cell size driven by low-contact-resistivity poly-Si diode. In *2009 Symposium on VLSI Technology* (pp. 24–25). IEEE 2009
61. N. Takaura, M. Kinoshita, M. Tai, T. Ohyanagi, K. Akita, T. Morikawa, A 4F2-cross-point phase change memory using nano-crystalline doped GeSbTe material. Jpn. J. Appl. Phys. **54** (4S), 04DD01 (2015)
62. S. Meister, S. Kim, J.J. Cha, H.S.P. Wong, Y. Cui, In situ transmission electron microscopy observation of nanostructural changes in phase-change memory. ACS Nano **5**(4), 2742–2748 (2011)
63. P.H. Bolivar, F. Merget, D.H. Kim, B. Hadam, H. Kurz, Lateral design for phase change random access memory cells with low-current consumption. Institute of Semiconductor Electronics, RWTH Aachen University Sommerfeldstr **24**, 52056 (2004)
64. F. Merget, D.H. Kim, P.H. Bolivar, H. Kurz, Lateral phase change random access memory cell design for low power operation. Microsyst. Technol. **13**(2), 169–172 (2007)
65. M.H. Lankhorst, B.W. Ketelaars, R.A.M. Wolters, Low-cost and nanoscale non-volatile memory concept for future silicon chips. Nat. Mater. **4**(4), 347–352 (2005)
66. Y.C. Chen, C.T. Rettner, S. Raoux, G.W. Burr, S.H. Chen, R.M. Shelby, M. Salinga, W.P. Risk, T.D. Happ, G.M. McClelland, M. Breitwisch, Ultra-thin phase-change bridge memory device using GeSb, in *International Electron Devices Meeting*, (2006), pp. 777–780
67. D. Krebs, S. Raoux, C.T. Rettner, G.W. Burr, M. Salinga, M. Wuttig, Threshold field of phase change memory materials measured using phase change bridge devices. Appl. Phys. Lett. **95** (8), 082101 (2009)
68. F. Xiong, A.D. Liao, D. Estrada, E. Pop, Low-power switching of phase-change materials with carbon nanotube electrodes. Science **332**(6029), 568–570 (2011)
69. J. Liang, R.G.D. Jeyasingh, H.Y. Chen, H.S.P. Wong, An ultra-low reset current cross-point phase change memory with carbon nanotube electrodes. IEEE Transactions on Electron Devices **59**(4), 1155–1163 (2012)
70. F. Xiong, M.H. Bae, Y. Dai, A.D. Liao, A. Behnam, E.A. Carrion, S. Hong, D. Ielmini, E. Pop, Self-aligned nanotube–nanowire phase change memory. Nano Lett. **13**(2), 464–469 (2013)
71. Y. Lu, S. Song, X. Shen, Z. Song, L. Wu, G. Wang, S. Dai, Low-power phase change memory with multilayer TiN/W nanostructure electrode. Applied Physics A **117**(4), 1933–1940 (2014)
72. Y. Lu, S. Song, Z. Song, L. Wu, A. He, Y. Gong, F. Rao, B. Liu, Superlattice-like electrode for low-power phase-change random access memory. Appl. Phys. Lett. **101**(11), 113104 (2012)
73. J.Y. Wu, M. Breitwisch, S. Kim, T.H Hsu, R. Cheek, P.Y. Du, J. Li, E.K. Lai, Y. Zhu, T.Y. Wang, H.Y. Cheng. A low power phase change memory using thermally confined TaN/TiN bottom electrode. In *Electron Devices Meeting (IEDM), 2011 I.E. International* (pp. 3–2). IEEE 2011
74. C. Ahn, S.W. Fong, Y. Kim, S. Lee, A. Sood, C.M. Neumann, M. Asheghi, K.E. Goodson, E. Pop, H.S.P. Wong, Energy-efficient phase-change memory with graphene as a thermal barrier. Nano Lett. **15**(10), 6809–6814 (2015)
75. T.C. Chong, L.P. Shi, R. Zhao, P.K. Tan, J.M. Li, H.K. Lee, X.S. Miao, A.Y. Du, C.H. Tung, Phase change random access memory cell with superlattice-like structure. Appl. Phys. Lett. **88** (12), 122114 (2006)

76. D. Loke, L. Shi, W. Wang, R. Zhao, H. Yang, L.T. Ng, K.G. Lim, T.C. Chong, Y.C. Yeo, Ultrafast switching in nanoscale phase-change random access memory with superlattice-like structures. Nanotechnology **22**(25), 254019 (2011)
77. R.E. Simpson, P. Fons, A.V. Kolobov, T. Fukaya, M. Krbal, T. Yagi, J. Tominaga, Interfacial phase-change memory. Nat. Nanotechnol. **6**(8), 501–505 (2011)
78. J. Momand, R. Wang, J.E. Boschker, M.A. Verheijen, R. Calarco, B.J. Kooi, Interface formation of two-and three-dimensionally bonded materials in the case of GeTe–Sb 2 Te 3 superlattices. Nano **7**(45), 19136–19143 (2015)
79. J. Tominaga, A.V. Kolobov, P. Fons, T. Nakano, S. Murakami. Ferroelectric order control of the Dirac-semimetal phase in GeTe-Sb2Te3 Superlattices. Adv. Mater. Interfaces, 1(1) (2014)
80. X. Yu, J. Robertson, Modeling of switching mechanism in GeSbTe chalcogenide superlattices. Sci Rep **5**, 12612 (2015)

Chapter 9
Phase-Change Memory Device Architecture

Fabio Pellizzer

9.1 Self-Heating Device Architectures

Self-heating PCM devices rely on the Joule heating generated inside the chalcogenide itself in order to increase the temperature and cause the phase transition.

Since the chalcogenide materials become quite conductive, both electrically and thermally, when their current density gets into the 1–10 MA/cm^2 range, amorphizing a self-heating device could involve higher current densities than a device comprising a heater. But the advantages of the heater-less solution are a simpler structure and the absence of all the wear-out mechanisms related to the heater material itself.

As described in [1], the higher current required by self-heating structures can be compensated by a dedicated optimization of the geometrical features of the device. For example, a smaller cross section available for the current flow will decrease the power requirements. Also an increased aspect ratio for a given cross section would obtain the same effect but could introduce critical issues for the process integration.

9.1.1 Pillar Structure

The pillar structure (Fig. 9.1) is the most compact and straightforward realization of a self-heating PCM cell, and it was proposed for the first time in [2], with a 75 nm diameter and 900 µA programming current.

F. Pellizzer (✉)
Micron Technology R&D, Boise, ID, USA
e-mail: fpellizz@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_9

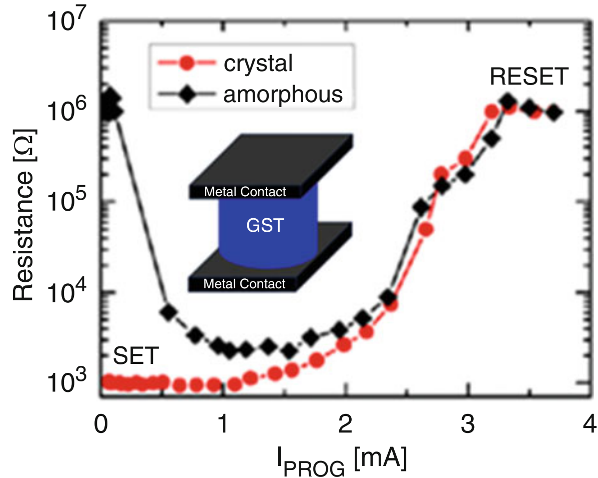


Fig. 9.1 R-I curves of a 45-nm-diameter cell, based on the vertical self-heating pillar architecture

The memory element is a fully confined parallelepiped of chalcogenide material, surrounded by electrodes on top and bottom and by sealing dielectrics on the sidewalls.

It is apparent that in this architecture, the critical choices are related to the patterning of the phase-change material.

The ideal solution is to etch the active materials (chalcogenide and electrodes) in one single step, using a dot-shaped pattern. But this approach could become very unpractical for very tight technology nodes (lateral dimension less than ~40 nm, with pitch lower than ~80 nm), and a two-step patterning would then be preferable, where the first direction is defined (Fig. 9.2) and then the second direction is defined after filling with dielectrics the trenches left by the first etch (Fig. 9.3).

In all the process synopses reported in this chapter, the memory element lands on a so-called WL Plug (word-line plug): this plug simply represents the connection to an underlying selecting device (MOSFET, Diode, etc.) that can be turned on with a dedicated word-line signal (WLs run in the x-direction) coupled with a dedicated bit-line signal (BLs run in the x-direction), in order to keep the description as general as possible.

In this case, the second etch becomes more critical because it has to go through different materials (filling dielectrics and chalcogenide/electrodes) at the same time.

The main limitation of this structure is the programming current that it can achieve without using a very high aspect ratio of the GST parallelepiped: in [1] they reported ~3 mA for an almost perfect cube of GST material with side 45 nm.

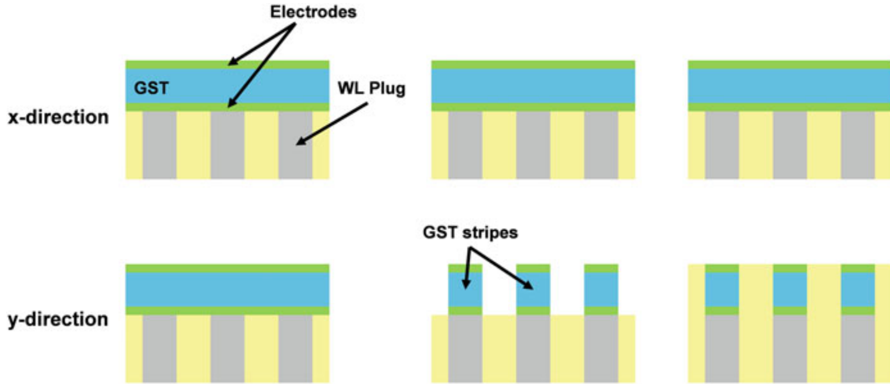


Fig. 9.2 Process synopsis for the definition of pillar structures in a first direction (e.g., along the word-lines)

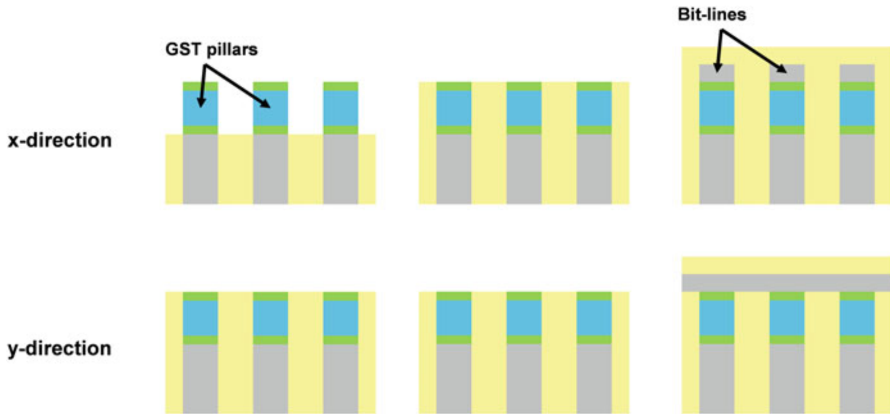


Fig. 9.3 Process synopsis for the definition of pillar structures in a second direction (e.g., along the bit-lines)

9.1.2 Self-Heating Wall Structure

In order to reduce the current I_{RESET} required by the pillar structure and, at the same time, maintain a very precise control of all the cell dimensions, the self-heating wall architecture can be used (Fig. 9.4).

The processing of this structure is described in Fig. 9.5, assuming that couples of adjacent bits (along the same bit-line) share the same "wall hole," in which the chalcogenide sidewalls are created.

The tight control on the smallest dimension is achieved through deposition of the chalcogenide layer on the sidewall of a dielectric material (silicon oxide or silicon nitride). Since the deposition has to be performed on a vertical surface and the wall holes can be quite narrow for a very dense array (the aspect ratio of the holes can be

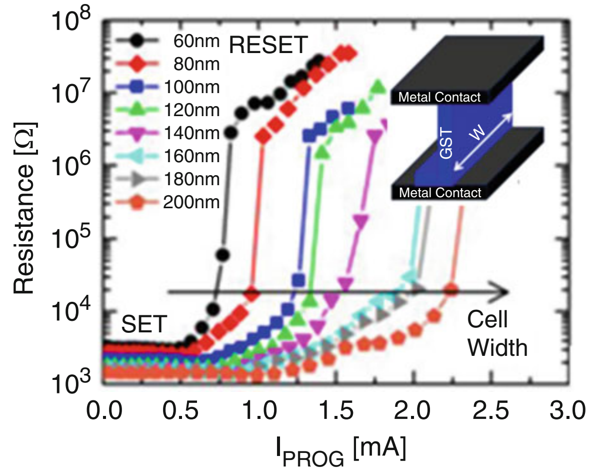


Fig. 9.4 R-I curves of cells with different widths, *W*, and based on vertical self-heating wall (SHW) architecture

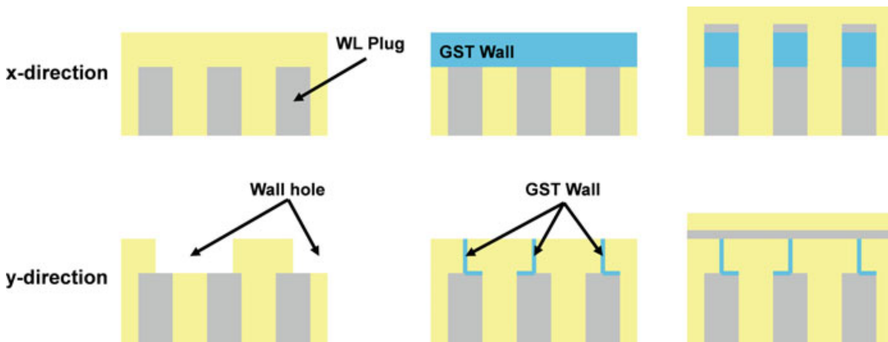


Fig. 9.5 Process synopsis for the self-heating wall structure integration

easily 1:1 or worse), a conformal deposition (e.g., chemical vapor deposition) of the phase-change material is preferred over a standard physical vapor deposition.

Another critical point of the process for the self-heating wall is the necessity of planarizing the surface and reexposing the top of the chalcogenide sidewall after the filling with dielectrics. In this case, a CMP would be the ideal solution, if compatible with the GST material.

9.1.3 Line Structure

The line structure [3] is created by patterning a thin layer of chalcogenide on top of a dielectric substrate in order to create a thin strip of phase-change material that gets to a wider and wider cross section until it is finally contacted by metal pads (Fig. 9.6).

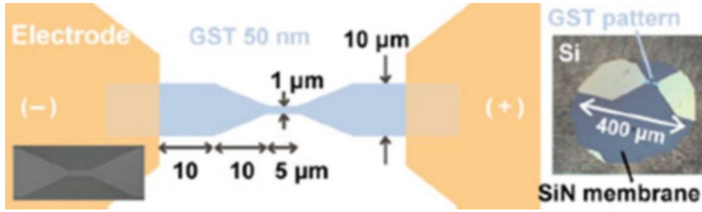


Fig. 9.6 SEM picture of an ultrathin phase-change bridge structure (© (2009) Applied Physics Letters, from [4])

As it is designed, the line structure is not optimized for very dense applications, and, in particular, the voltage drop and the power dissipated in the fan-out regions are not sustainable in a real product.

This structure is nevertheless very useful for basic studies related to the movement of the elements inside the chalcogenide material due to electrothermal and mechanical driving forces [5].

9.1.4 Bridge Structure

The bridge structure [6] represents an evolution of the line toward a more compact device to be used in a real memory architecture.

In this case, two metallic electrodes are separated by a narrow dielectric layer, and a thin chalcogenide strip connects the two electrodes by crossing the top surface of the dielectric layer (Fig. 9.7).

Without the additional losses in the fan-out region of the bridge, this device was able to reach an I_{RESET} of 90 μA and reset voltage of 1.5 V, featuring a total volume of 20 nm × 50 nm × 3 nm (3000 nm^3).

Even if this structure is indeed much more compact than the line, it is still not the best solution for a very dense array of phase-cell memory cells, because of its lateral extension that prevents the minimum pitch design (also referred to as "4·F^2," i.e., 2·F × 2·F, where F is the minimum half-pitch achievable by a given technology node) typical of other structures which extend only in the vertical direction. Nevertheless, its simplicity, the tight control of the GST cross section (thickness and etching), and the planar surface for the chalcogenide alloy deposition make it a possible choice for less dense applications.

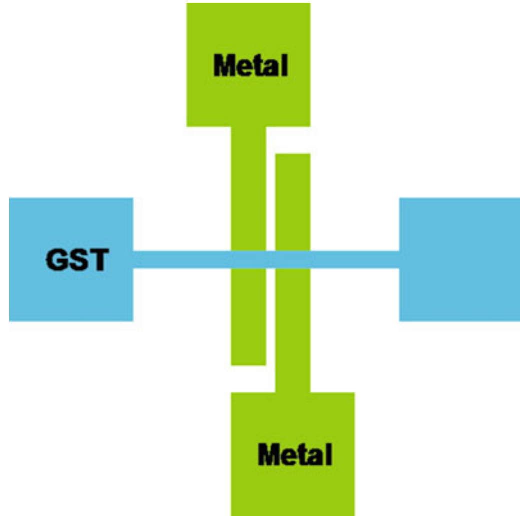


Fig. 9.7 Plan view schematic for a bridge structure, featuring a thin strip of chalcogenide material running on top of dielectric between two metallic lines

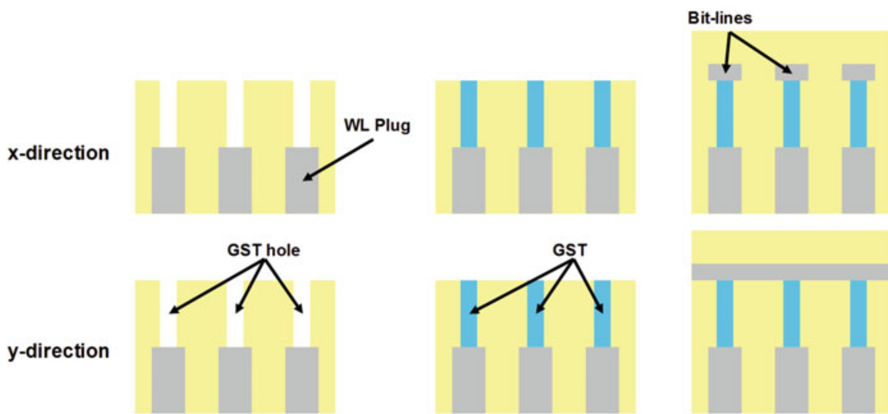


Fig. 9.8 Process synopsis for the self-heating via structure integration

9.1.5 Via Structure

The pillar structure discussed above has a natural counterpart in the via structure, in which the GST material fills a previously created hole in a dielectric, as described in Fig. 9.8.

As opposed to the pillar structure, in this case the lateral side of the active chalcogenide material has not to be etched: the only critical modification of the GST could come from the etch-back (or the polishing) needed in order to remove it from the top of the dielectric.

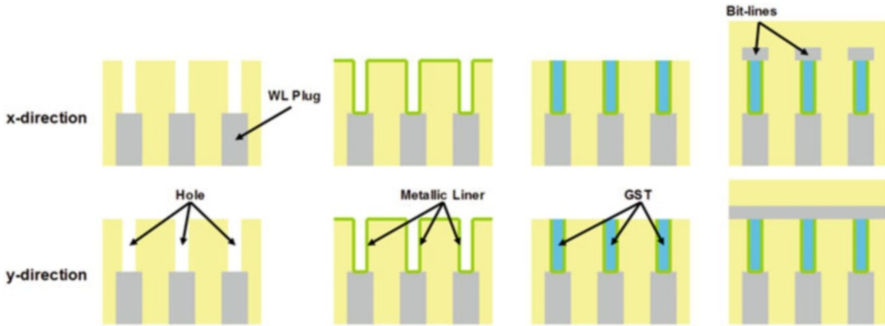


Fig. 9.9 Process synopsis for the metallic liner structure integration

On the other side, a very conformal deposition of the phase-change alloy is required if the via has a high aspect ratio, as one would try to pursue in order to minimize the programming current.

9.1.6 Metallic Liner Structure

An interesting evolution of the via structure is represented by the metallic liner structure, discussed in [7] and illustrated in Fig. 9.9.

In this architecture the metallic liner has two functions:

- To mitigate the drift of the amorphous resistance, due to the controlled resistance in parallel with the GST resistance
- To improve the reliability of the cell, by preventing void formation during electrical operation

The mitigation of the amorphous phase drift is an important factor in order to enable multilevel cell (MLC) capability for PCM memory, since it would allow a better separation among the intermediate (and partially amorphous) states during the operating life of the device.

9.2 Built-In Heater Device Architectures

Self-heating structures can count only on geometrical optimization in order to keep I_{RESET} in a workable range [1], and for this reason many efforts have been spent in order to introduce a dedicated element, called the "heater," inside the PCM cells to generate the heating necessary for the phase transition.

The introduction of this new element opened a new field of research to find the heater materials that were compatible with the phase-change chalcogenide alloys, because such materials need to have:

- The right resistivity, i.e., 1–10 mOhm·cm, in order to provide a good generation of power
- Good compatibility (no cross-contamination during process and operation) with the phase-change alloy
- Good adhesion with the chalcogenide material, in order to prevent interfacial delamination during operation
- Good endurance (i.e., stable resistivity) when subject to high current density ($>$10 MA/cm^2) for many cycles ($>10^6$)

9.2.1 Pore Structure

The pore structure was indeed the first structure [8] used to create chalcogenide-based electron devices, because of the ease of its realization.

Figure 9.10 shows the SEM view of such structure in one of its first realizations, and Fig. 9.11 illustrates the cross section of the structure and its process synopsis.

Building the pore cell does not require, in principle, any chemical-mechanical polishing (CMP) step, and this is one of the reasons why it was preferred at the dawn of process integration. Moreover, a very relaxed patterning of the chalcogenide material can be used, since the active area is determined by the area of the hole opened in the dielectric material (previously deposited on top of the bottom electrode material).

More recently an aggressive patterning of the pore was proposed in [9], in order to create a very small contact area between the chalcogenide material and the underlying heater, as shown in Fig. 9.12.

The programming current, ~250 μA, reported for this structure is indeed competitive for a similar technology node, but the manufacturability of such sublithographic approach (through poly-Si spacers) could be questionable, and with very aggressive scaling (i.e., half-pitch below 40 nm), it would become unpractical.

9.2.2 Matchstick Structure

The matchstick structure (Fig. 9.13) represented one of the first attempts to reduce the programming current of a phase-change cell with a built-in heater after the CMP became a widely used and manufacturable process [10].

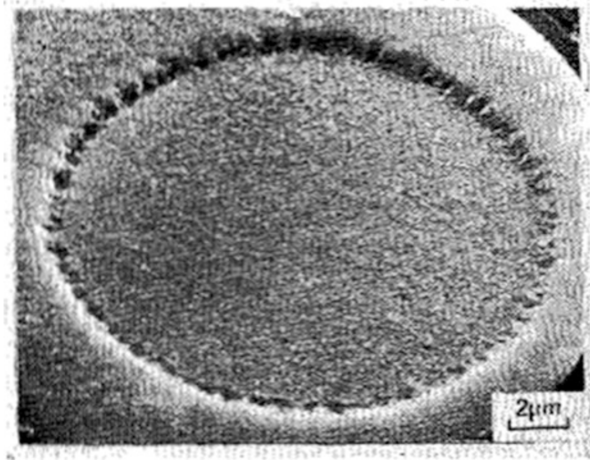


Fig. 9.10 SEM picture of a phase-change memory material in a 20-μm-diameter pore. The top electrode was removed by etching after electrical testing (© (1973) IEEE, from [8])

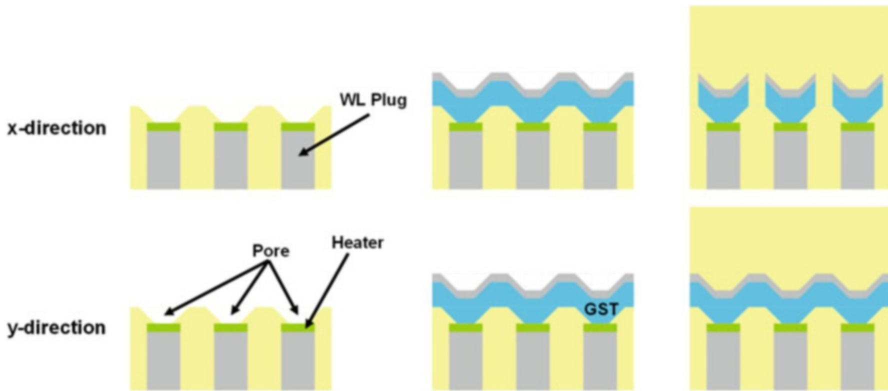


Fig. 9.11 Process synopsis for the pore structure integration

Figure 9.14 shows a possible process synopsis for this structure, focusing on the steps required for the creation of the heater and the subsequent definition of GST bit-lines.

The realization of the matchstick requires a very conformal metallic (or semimetallic) layer in order to fill a high aspect ratio hole that defines the critical dimension for the contact area between the matchstick and the chalcogenide alloy.

Moreover, such material has to be friendly for CMP and has to be resilient to oxidation and/or easy to clean, in order to minimize its interface resistance with the phase-change material.

In the past, there have been proposals for engineering the interface resistance between heater and chalcogenide to better control/reduce the programming current [11].

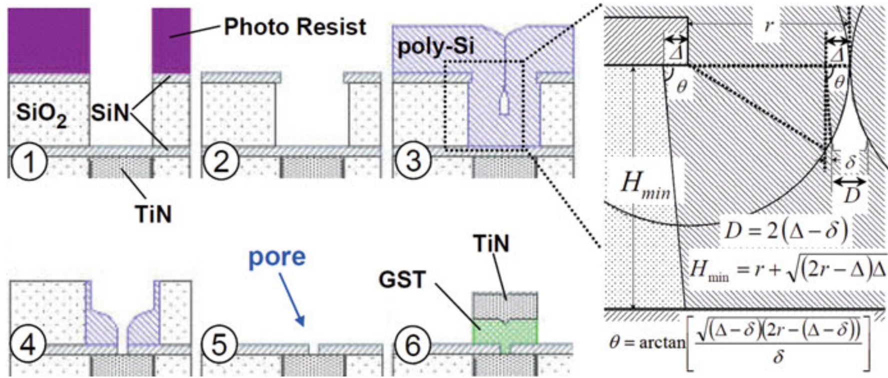


Fig. 9.12 Pore memory element process modules and keyhole size parameters (© (2007) IEEE, from [9])

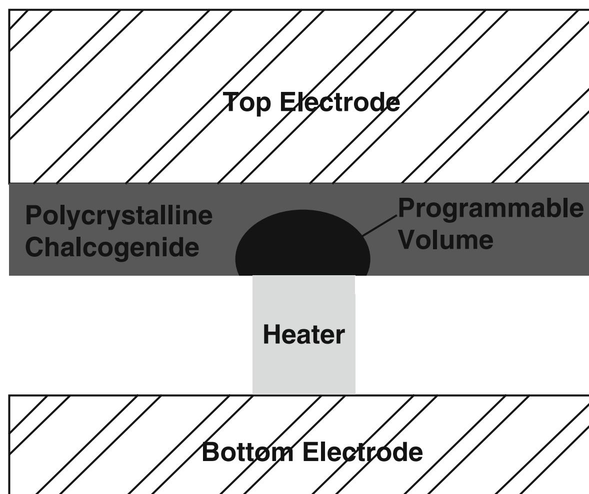


Fig. 9.13 Schematic cross section of the memory element (© (2001) IEEE, from [10])

As illustrated in Fig. 9.15, these interfacial layers can also improve the adhesion of the chalcogenide material on the surface coplanar with the heater top. In fact, the top surface of the heater is laterally surrounded by dielectric materials (e.g., silicon oxide, silicon nitride, etc.), and careful process optimization is required in order to promote a good adhesion between GST and such layers.

Even if these interfacial layers can be very effective in reducing the required I_{RESET} for a given technology node (the reset current density is ~1 MA/cm^2 in [11], compared to the typical 10–100 MA/cm^2 of other PCM structures), their endurance is typically a weak spot for such structures, because it is not easy to find a dielectric material able to sustain the current densities required by the amorphization of the phase-change alloy for an extended period of time (e.g., one million cycles with

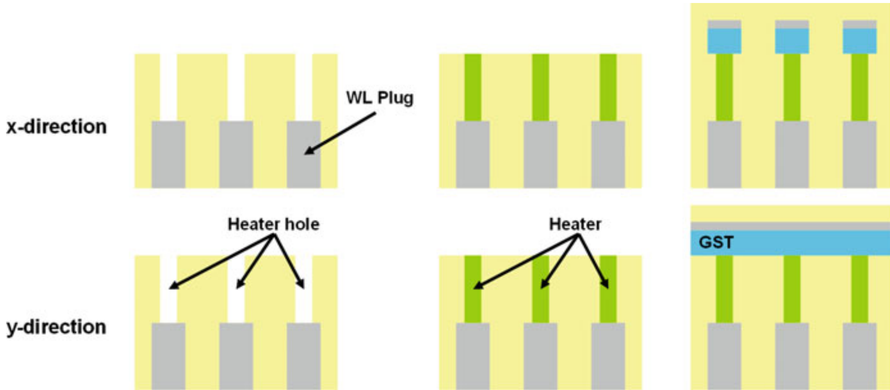


Fig. 9.14 Process synopsis for matchstick structure integration

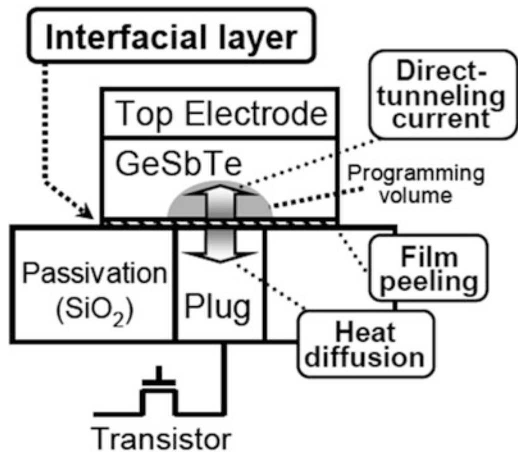


Fig. 9.15 Schematic cross section of phase-change memory cell using interfacial layer to suppress film peeling of GST and reduce heat diffusion from GST to plug. Joule heat may be generated by direct tunneling current through the interfacial layer (© (2006) IEEE, from [11])

t_{RESET} equal to 50 ns gives a charge fluence through the dielectric equal to $5 \cdot 10^4$ C/cm^2).

In some cases, the interfacial layers can just be used as adhesion promoters (Fig. 9.16).

As reported in [12], metal layers typically provide better adhesion for chalcogenide materials, but when they are in close contact with the region that undergoes the phase transition, they must have a very low reactivity with GST at high temperatures in order to minimize intermixing and poisoning during device operation.

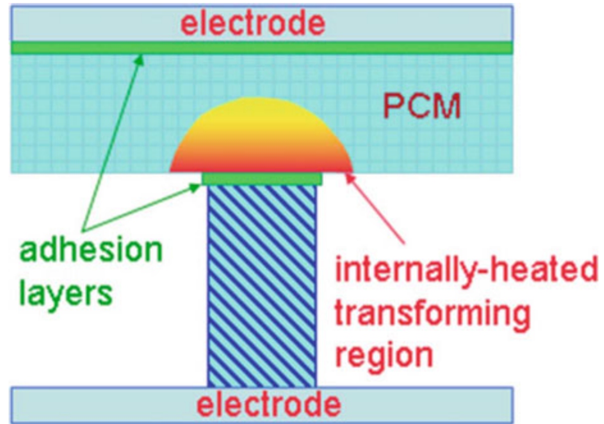


Fig. 9.16. Schematic cross section of a phase-change memory cell. The critical adhesion layer is in direct contact with the transforming region (© (2007) AIP, from [12])

9.2.3 Ring Structure

The ring structure was proposed by [13] and [14] in order to reduce the programming current of the matchstick structure with a minimal architectural difference: instead of depositing the heater material to fill the high aspect ratio hole, only a thin layer of metal (or semimetal) was deposited, and then the rest of the hole was filled with a dielectric material (Fig. 9.17).

By virtue of the smaller contact area, a reduction of ~10% on I_{RESET} was reported, in agreement with simulation results (Fig. 9.18).

It is worth noting that in this case, even if the contact area was reduced more than 50%, the reduction of programming current was not following the same scaling rule. This effect underlines an important factor to be considered in the design of any phase-change structure: in order to reduce effectively the programming current, we have to reduce both the top and the bottom equivalent thermal resistance.

If one of those two thermal resistances is much lower than the other, the gain of lowering the other one will be greatly reduced, as it happens in this ring structure where the top thermal resistance is the lower and limiting one. More details on the cell electrothermal behavior can be found in Chap. 3.

A critical integration issue could be the interaction between the filling dielectric and the heater. This dielectric should be, in principle, very conformal in order to close perfectly the structure. But even if some cavity would remain that would not necessarily cause a failure of the structure, unless the CMP before chalcogenide deposition does re-expose the cavity and lead to a defect. Moreover, the dielectric has to be compatible (i.e., not damaging and/or not oxidizing) for the heater material.

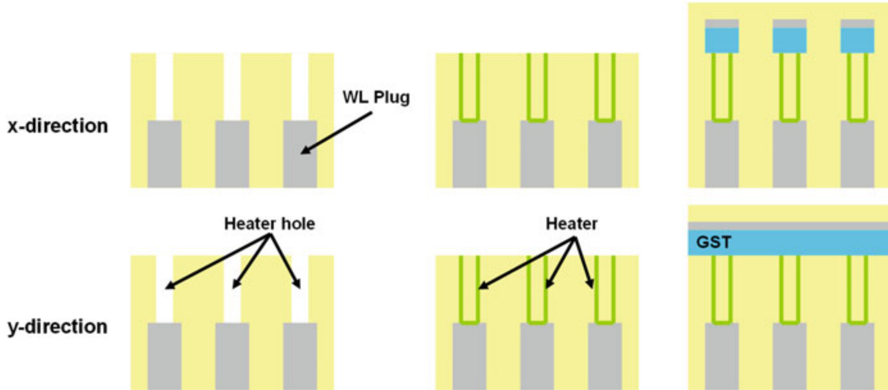


Fig. 9.17 Process synopsis for ring structure integration

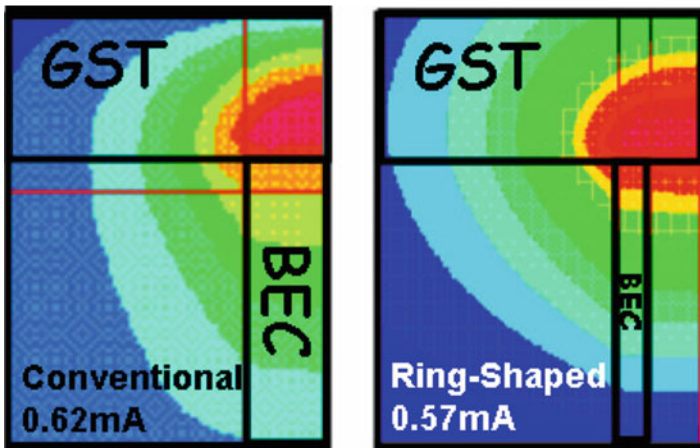


Fig. 9.18 Simulation results showing the 50 uA writing current reduction of the ring-shaped contact (© (2005) IEEE, from [13])

9.2.4 µTrench Structure

An alternative way of reducing the contact area between the phase-change material and the heater is represented by the µTrench ("micro-trench") structure.

The first realization of the µTrench was reported in 180 nm technology [15] and was then scaled down to the 90 nm node [16].

The basic concept consists in creating a contact region limited, in one direction, by the thickness of the deposited heater and, in the other direction, by the opening created in a dielectric layer (e.g., silicon nitride), as illustrated in Fig. 9.19.

The main issue and non-ideality of the µTrench structure resides in the difficulty of obtaining a perfectly flat surface between the chalcogenide material and the

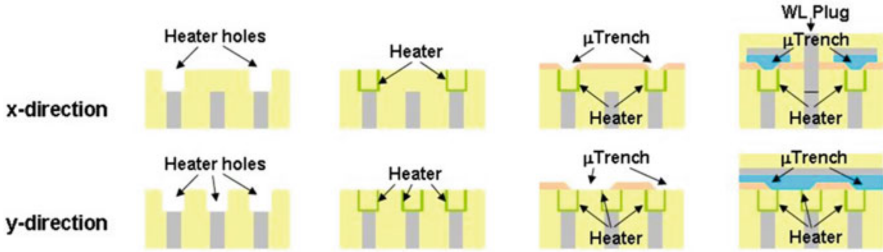


Fig. 9.19 Process synopsis for the standard µTrench structure (© (2007) IEEE, from [17])

vertical heater, because there is always some differential in the etch selectivity between the dielectrics surrounding the heater and the heater itself.

Nevertheless, this structure demonstrated a reset operation with 700 µA at the 90 nm technology node, and a self-aligned evolution was shortly developed [17], and it is shown in Fig. 9.20.

The self-aligned architecture was also optimized for the most important geometrical parameters (heater thickness, heater height, and GST thickness) in order to achieve a reset current of 450 µA at the same technology node.

Figure 9.21 shows the qualitative response of the programming current to each of the three geometrical factors, in order to illustrate the directions pursued for the optimization.

9.2.5 Wall Structure

In order to improve the controllability of the surface chalcogenide heater and to complete the self-alignment development of the µTrench, a novel architecture, the wall, was introduced in [18].

As illustrated in Fig. 9.22, in the wall structure, the top contact surface of the vertical heater is finished with a chemical mechanical polishing (CMP) that is able to provide a more planar morphology, compared to a dry etching.

The wall structure is quite compact, and its manufacturability was demonstrated with a 1Gb PCM product operating at 1.8 V and exploiting the low I_{RESET}, 200 µA, achieved at the 45 nm technology node [19].

This result was possible because the wall structure does likely give an optimal trade-off among several figures of merit required for VLSI devices.

The radar chart in Fig. 9.23 demonstrates the better trade-off given by the wall heater, which for a small disadvantage in efficiency (ratio between power used for phase-change and power generated inside the structure) provides very small process and integration complexity and several optimization knobs.

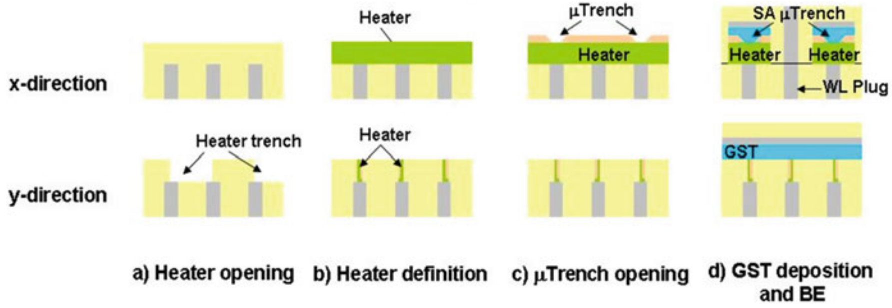


Fig. 9.20 Process synopsis for the self-aligned µTrench structure integration (© (2007) IEEE, from [17])

Parameter	Variation	I_{prog}	V_{prog}	R_{set}
Heater thickness	↑	↑	~	↓
	↓	↓	~	↑
Heater height	↑	↓	↑	↑
	↓	↑	↓	↓
GST thickness	↑	↓	↑	↑
	↓	↑	↓	↓

Fig. 9.21 Expected trends (higher ↑, lower ↓, almost equal ~) for the main electrical parameters as a function of the geometrical variations of some cell structure parameters (© (2006) IEEE, from [17])

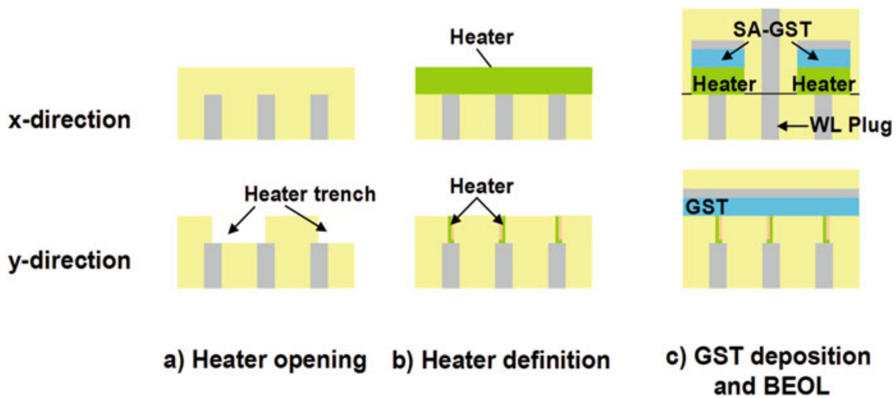


Fig. 9.22 Process synopsis for the wall structure integration: (**a**) Heater opening, (**b**) Heater definition and (**c**) GST deposition and BEOL (© (2014) IEEE, from [18])

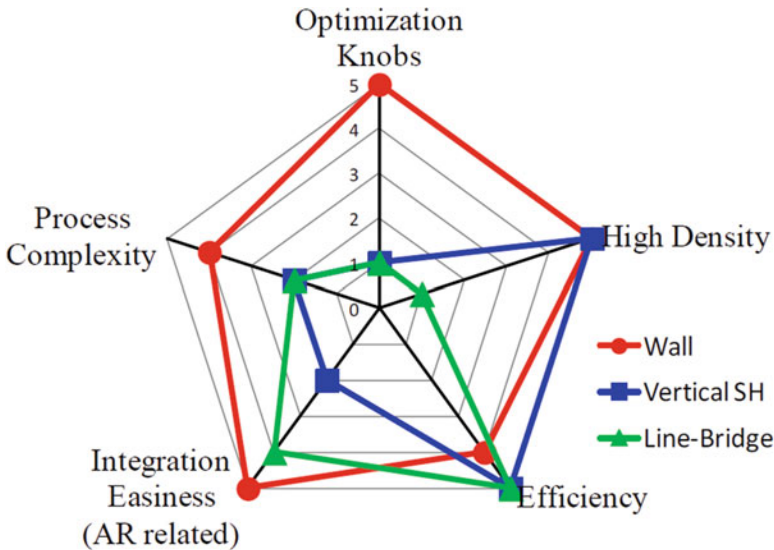


Fig. 9.23 Summary of strengths and drawbacks for three architectures: wall heater, vertical self-heating (either pillar or wall), and line bridge (© (2014) IEEE, from [1])

9.2.6 Dash-Confined Structure

An interesting development of structures based on a vertical heater wall has been shown by [20]. They demonstrated a dash-confined cell (Figs. 9.24 and 9.25), capable of reaching I_{RESET} as low as 160 μA, with a minimum chalcogenide thickness of 7.5 nm.

The innovative aspect of this structure is the introduction of a replacement flow for the control of the phase-change volume.

In this case the chalcogenide material replaces a previously deposited BEC (bottom electrode contact) metallic layer inside a very small slot created with a partial wet etch removal (recess) of the BEC itself.

In order to have a tight control on the morphology and characteristics of the PCM alloy, a very conformal deposition is needed, since the aperture available for the filling is 7.5 nm wide and 30 nm deep. Indeed, this kind of replacement flows typically requires chemical vapor deposition (CVD) of the phase-change layer.

Moreover, the control and the cleaning of the recession are critical factors in the performance of the device, because they determine the quality of the interface between the phase-change alloy and the recessed heater, and we already illustrated as this interface influences I_{RESET} and cycling endurance of PCM.

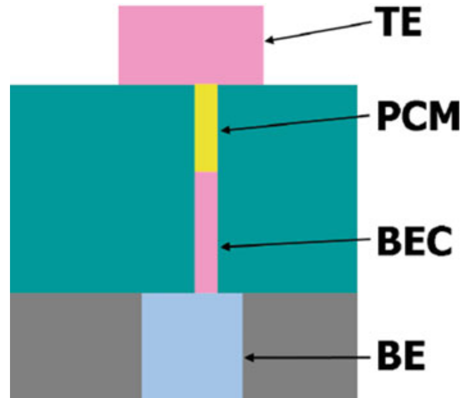


Fig. 9.24 Schematic diagram of dash-confined cell structure (© (2008) IEEE, from [20])

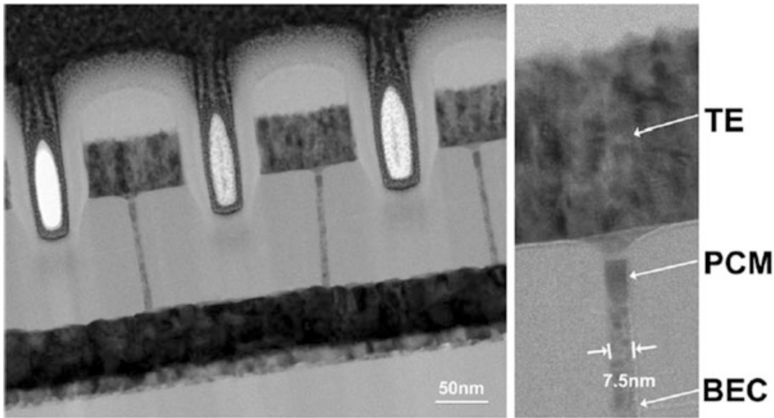


Fig. 9.25 TEM images of dash-type confined cell structure. The width of PCM in the contact is approximately 7.5 nm (© (2008) IEEE, from [20])

9.2.7 Cross-Spacer Structure

The extreme evolution of architectures based on a vertical heater lamina is represented by the cross-spacer structure, reported by [21].

Figure 9.26 shows the schematic diagrams of this cell, and we can appreciate that both the heater and the phase-change material are deposited on sidewalls, thus merging the approaches of the self-heating wall and of the μTrench/wall architectures.

Even if this approach opened up the possibility for a dramatic programming current reduction on less recent technology nodes, it does apparently run out of steam when the half-pitch becomes very small, <30 nm, because the opening for heater and GST deposition, the dimensions of the active sidewalls, and the

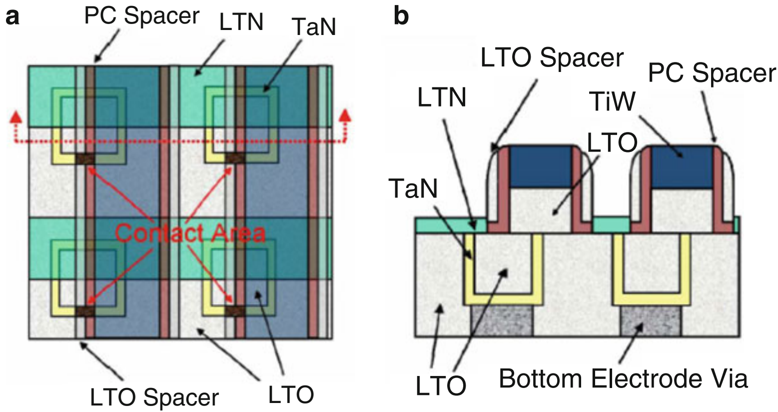


Fig. 9.26 Schematic diagrams of cross-spacer cell structure: (**a**) top view and (**b**) side view (© (2007) IEEE, from [21])

extension of the spacers start to be quite comparable with the minimum dimensions already achievable with the technology.

9.3 Remote Heater Device Architectures

A heater in direct contact with the chalcogenide material and not requiring additional connections is the preferred choice in order to maintain a compact device structure, suitable for aggressive memory cell footprint.

But a heater physically separated from the chalcogenide material offers the possibility for independent optimization of material parameters. For example, the optimization of the heater properties does not have to take into account the compatibility between the heater itself and the chalcogenide material when they are separated by a dielectric layer.

Moreover, these structures open the opportunity for a characterization of the thermal behavior of the phase-change material without convolving it with the flux of electrical current through the material itself.

9.3.1 Micro-Heater Structure

A good example of such structure is reported in [22] and [23], where the remote heater is used in order to induce the phase-change in a chalcogenide material separated by a dielectric barrier, which prevents a direct electrical interaction between the two parts of the system (Fig. 9.27).

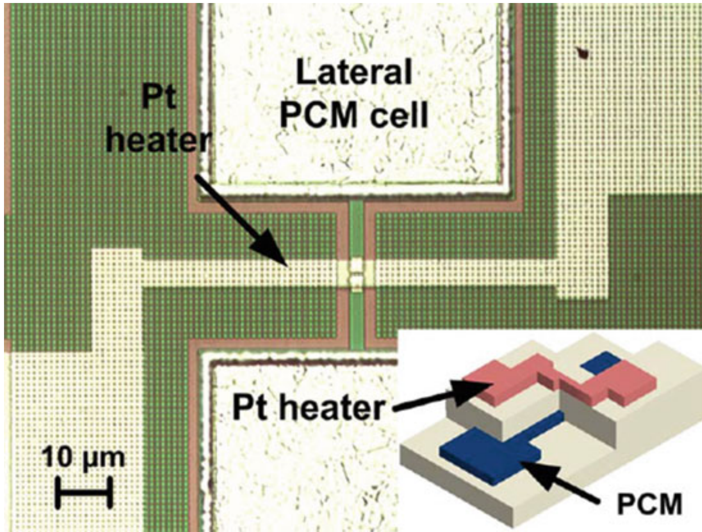


Fig. 9.27 Microscope image of an MTS. Pt heater is integrated on top of the lateral PCM cell (three-dimensional figure inset) (© (2015) IEEE, from [22])

This structure is also referred as a micro-heater and can be engineered and simulated in order to provide accurate estimation of the temperatures reached during operation [24].

Typically, the metallic layers used as a heater need to be able to sustain quite high temperatures, because they need to generate enough power to heat the chalcogenide above melting (>600 °C), and preferred solutions include tungsten [22] and platinum [23].

An extreme evolution of these structures based on a micro-heater can be found in [25], where the carbon nanotube (CNT) heater with sub-5 nm diameter is back in contact with the phase-change alloy, as illustrated in Fig. 9.28.

The programming current reported for this structure is low, because crystallization of the GST surrounding the nanotube starts at ~25 μA (even if the voltage is quite high, ~ 15 V, due to the long and resistive nanotube), but the direct axial contact between heater and GST prevents the possibility of re-amorphizing the crystallized volume. Nevertheless, this demonstration opens up the possibility of developing phase-change memory cells with CNT heaters and ultralow switching energy, possibly using a different architecture.

9.4 Summary

In this chapter we have discussed the options and the evolution of device architectures for phase-change memories.

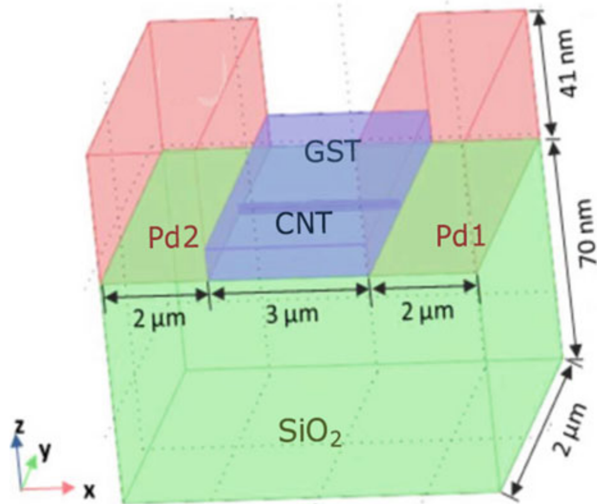


Fig. 9.28 Carbon nanotube heater structure used in the 3D simulation, which mimics the experimental test structures (© (2009) AIP, from [25])

After this review, it is clear that the actual choice of the optimal cell structure depends on many boundary conditions:

- The area available for the cell (i.e., the density of the actual device)
- The driving current delivered by the selecting device
- The number of masks available for the cell integration
- The power consumption
- The technology node
- The required resistance of the storage element

For example, the choice of using a heater inside the cell is likely optimal if a low-current, high-efficiency, and compact solution is needed. And an architecture as the self-aligned wall heater can provide a demonstrated manufacturable solution.

As the phase-change memory technology enters into the storage class memory contest [26] as extensively discussed in Chap. 11, some of these cell architectures will be challenged and improved in order to deliver the required performances (speed, endurance, and cost).

References

1. M. Boniardi, A. Redaelli, C. Cupeta, F. Pellizzer, L. Crespi, G. D'Arrigo, A.L. Lacaita, G. Servalli, Optimization metrics for phase change memory (PCM) cell architectures. in *International Electron Devices Meeting*, 2014, pp. 681–684
2. T. Happ, M. Breitwisch, A. Schrott, J. Philipp, M. Lee, R. Cheek, T. Nirschl, M. Lamorey, C. Ho, S. Chen, C. Chen, E. Joseph, S. Zaidi, G. Burr, B. Yee, Y. Chen, S. Raoux, H. Lung, R. Bergmann, C. Lam, Novel one-mask self-heating pillar phase change memory. in *IEEE Symposium on VLSI Technology Digest of Technical Papers*, 2006 pp.120–121

3. M.H.R. Lankhorst, B.W.S.M.M. Ketelaars, R.A.M. Wolters, Low-cost and nanoscale non-volatile memory concept for future silicon chips. Nat. Mater. **4**(4), 347–352 (2005)
4. D. Kang, D. Lee, H.-M. Kim, S.-W. Nam, M.-H. Kwon, K.-B. Kim, Analysis of the electric field induced elemental separation of $Ge_2Sb_2Te_5$ by transmission electron microscopy. Appl. Phys. Lett. **95**(011904), 1–3 (2009)
5. L. Crespi, A. L. Lacaita, M. Boniardi, E. Varesi, A. Ghetti, A. Redaelli, G. D'Arrigo, Modeling of atomic migration phenomena in phase change memory devices. in *IEEE International Memory Workshop*, 2015 pp. 1–4
6. Y.C. Chen, C.T. Rettner, S. Raoux, G.W. Burr, S.H. Chen, R.M. Shelby, M. Salinga, W.P. Risk, T.D. Happ, G.M. McClelland, M. Breitwisch, A. Schrott, J.B. Philipp, M.H. Lee, R. Cheek, T. Nirschl, M. Lamorey, C.F. Chen, E. Joseph, S. Zaidi, B. Yee, H.L. Lung, R. Bergmann, C. Lam, Ultra-thin phase-change bridge memory device using GeSb. in *International Electron Devices Meeting*, 2006, pp. 1–4
7. W. Kim, M. BrightSky, T. Masuda, N. Sosa, S. Kim, R. Bruce, F. Carta, G. Fraczak, H.Y. Cheng, A. Ray, Y. Zhu, H.L. Lung, K. Suu, C. Lam, ALD-based confined PCM with a metallic liner toward unlimited endurance. in *International Electron Devices Meeting*, 2016 pp. 83–86
8. S.R. Ovshinsky, H. Fritzsche, Amorphous semiconductors for switching, memory, and imaging applications. IEEE Trans. Electron Devices **20**, 91–105 (1973)
9. M. Breitwisch, T. Nirschl, C.F. Chen, Y. Zhu, M.H. Lee, M. Lamorey, G.W. Burr, E. Joseph, A. Schrott, J.B. Philipp, R. Cheek, T.D. Happ, S.H. Chen, S. Zaidi, P. Flaitz, J. Bruley, R. Dasaka, B. Rajendran, S. Rossnagel, M. Yang, Y.C. Chen, R. Bergmann, H.L. Lung, C. Lam, Novel lithography-independent pore phase change memory. in *IEEE Symposium on VLSI Technology Digest of Technical Papers*, 2007 pp. 100–101
10. S. Lai, T. Lowrey, OUM–A 180 nm nonvolatile memory cell element technology for standalone and embedded applications. in *International Electron Devices Meeting*, 2001 pp. 803–806
11. Y. Matsui, K. Kurotsuchi, O. Tonomura, T. Morikawa, M. Kinoshita, Y. Fujisaki, N. Matsuzaki, S. Hanzawa, M. Terao, N. Takaura, H. Moriya, T. Iwasaki, M. Moniwa, T. Koga, Ta_2O_5 interfacial layer between GST and W plug enabling low power operation of phase change memories. in *International Electron Devices Meeting*, 2006 pp. 769–772
12. C. Cabral Jr., K.N. Chen, L. Krusin-Elbaum, V. Deline, Irreversible modification of Ge2Sb2Te5Ge2Sb2Te5 phase change material by nanometer-thin Ti adhesion layers in a device-compatible stack. Appl. Phys. Lett. **90**, 051908 (2007)
13. S.J. Ahn, Y.N. Hwang, Y.J. Song, S.H. Lee, S.Y. Lee, J.H. Park, C.W. Jeong, K.C. Ryoo, J.M. Shin, J.H. Park, Y. Fai, J.H. Oh, G.H. Koh, G.T. Jeong, S.H. Joo, S.H. Choi, Y.H. Son, J.C. Shin, Y.T. Kim, H.S. Jeong and Kinam Kim. Highly reliable 50nm contact cell technology for 256Mb PRAM. in *IEEE Symposium on VLSI Technology Digest of Technical Papers*, 2005 pp. 98–99
14. J.H. Oh, J.H. Park; Y.S. Lim; H.S. Lim; Y.T. Oh; J.S. Kim; J.M. Shin; J.H. Park; Y.J. Song; K.C. Ryoo; D.W. Lim; S.S. Park; J.I. Kim; J.H. Kim; J. Yu; F. Yeung; C.W. Jeong; J.H. Kong; D.H. Kang; G.H. Koh; G.T. Jeong; H.S. Jeong; K. Kim, Full integration of highly manufacturable 512Mb PRAM based on 90nm. in *IEEE International Electron Devices Meeting*, 2006 pp. 1–4
15. F. Pellizzer, A. Pirovano, F. Ottogalli, M. Magistretti, M. Scaravaggi, P. Zuliani, M. Tosi, A. Benvenuti, P. Besana, S. Cadeo, T. Marangon, R. Morandi, R. Piva, A. Spandre, R. Zonca, A. Modelli, E. Varesi, T. Lowrey, A. Lacaita, G. Casagrande, P. Cappelletti, R. Bez, Novel μTrench phase-change memory cell for embedded and stand-alone NVM applications. Proc. Symp. VLSI Technol., 18–19 (2004)
16. F. Pellizzer, A. Benvenuti, B. Gleixner, Y. Kim, B. Johnson, M. Magistretti, T. Marangon, A. Pirovano, R. Bez, G. Atwood *A 90nm Phase Change Memory Technology for Stand-Alone NVM Applications*. in *Proceedings Symposium on VLSI Technology* (2006)

17. A. Pirovano, F. Pellizzer, I. Tortorelli, R. Harrigan, M. Magistretti, P. Petruzza, E. Varesi, D. Erbetta, T. Marangon, F. Bedeschi, R. Fackenthal, G. Atwood, R. Bez. Self-aligned μTrench phase-change memory cell architecture for 90nm technology and beyond. in *Proceedings of ESSDERC*, 2007 pp. 222–225
18. G. Servalli, A 45nm generation phase change memory technology. in *Proc IEEE International Electron Devices Meeting*, 2009 pp. 1–4
19. C. Villa, D. Mills, G. Barkley, H. Giduturi, S. Schippers, D. Vimercati, A 45nm 1Gb 1.8V phase-change memory. ISSCC **53**, 270–271 (2010)
20. D.H. Im, J. I. Lee, S.L. Cho, H.G. An, D.H. Kim, I.S. Kim, H. Park, D.H. Ahn, H. Horii, S.O. Park, U. I. Chung, and J.T. Moon, A unified 7.5nm dash-type confined cell for high performance PRAM device. in, *IEEE International Electron Devices Meeting*, 2008 pp. 1–4
21. W. S. Chen, C.M. Lee, D.S. Chao, Y.C. Chen, F. Chen, C.W. Chen, P.H. Yen, M.J. Chen, W.H. Wang, T.C. Hsiao, J.T. Yeh, S.H. Chiou, M.Y. Liu, T.C. Wang, L.L. Chein, C.M. Huang, N.T. Shih, L.S. Tu, D. Huang, T.H. Yu, M.J. Kao, M.-J. Tsai, A novel cross-spacer phase change memory with ultra-small lithography independent contact area. in, *IEEE International Electron Devices Meeting*, 2007 pp. 319–322
22. S. Kim, B. Lee, M. Asheghi, F. Hurkx, J.P. Reifenberg, K.E. Goodson, H.S.P. Wong, Resistance and threshold switching voltage drift behavior in phase-change memory and their temperature dependence at microsecond time scales studied using a micro-thermal stage. IEEE Transactions on Electron Devices **58**, 584–592 (2011)
23. R. Singh, G. Slovin, M. Xu, A. Khairi, S. Kundu, T.E. Schlesinger, J.A. Bain, J. Paramesh, A 3/5 GHz reconfigurable CMOS low-noise amplifier integrated with a four-terminal phase-change RF switch. in *IEEE International Electron Devices Meeting*, 2015 pp. 636–639
24. M. Xu, G. Slovin, J. Paramesh, T.E. Schlesinger, J.A. Bain, Thermometry of a high temperature high speed micro heater. Rev. Sci. Instrum. **87**(2), 024904 (2016)
25. F. Xiong, A. Liao, E. Pop, Inducing chalcogenide phase change with ultra-narrow carbon nanotube heaters. Appl. Phys. Lett. **95**, 243103 (2009)
26. C.H. Lam, Storage class memory. in *IEEE international conference on solid-state and integrated circuit technology*, 2010 pp. 1080–1083

Chapter 10
PCM Array Architecture and Management

Corrado Villa

10.1 Array Definition

The actual implementation of a diode selected PCM array is strongly connected with the specification targets. There are several parameters like power or read and program latency that vary substantially with the architecture choices taken [1]. In any case, it is very likely that a high density array has to be fragmented in several sub-arrays in order to control leakage, power, and speed at the cost of some die size adder.

Several solutions are possible, and Fig. 10.1 shows a possible implementation of a PCM array core.

The core is made of several tiles, each comprising a column decoder, a row decoder, and a local sense amplifier. Tiles are grouped in partitions, which include a partition driver and some high-voltage switches that connect the partition to a subset of several power rails running across the die; partitions can be grouped further in planes, with each plane including a set of sense amplifiers and program loads.

Tiles are made of word lines and bit lines. Word lines are selecting one or more cells in parallel during read and write operations. Bit lines are always singly connected to a sense amp for reading or to a program load for writing. The cell activation through a diode selector requires the forward biasing of the diode itself, while all the other diodes corresponding to unselected cells must be either reverse biased or biased at a voltage that is well below a certain threshold; this is to guarantee that the current flowing in the deselected cell is (1) not causing any read or program disturb on the deselected cell itself and (2) not inducing a wrong sensing of the selected cell because of the excessive leakage (Fig. 10.2).

C. Villa (✉)
Micron, Vimercate, MB, Italy
e-mail: corradovilla@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_10

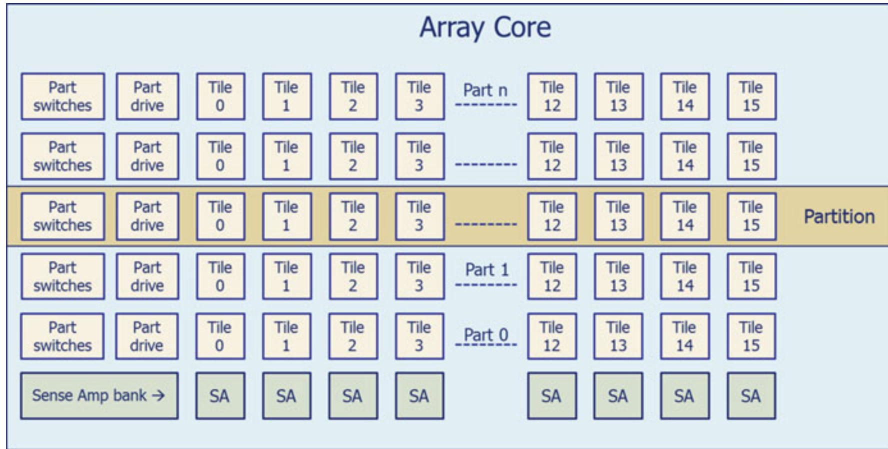


Fig. 10.1 A PCM array implementation. "Tiles" are grouped horizontally to realize a "partition," and a group of partitions (bank or plane) share a sense amplifier bank

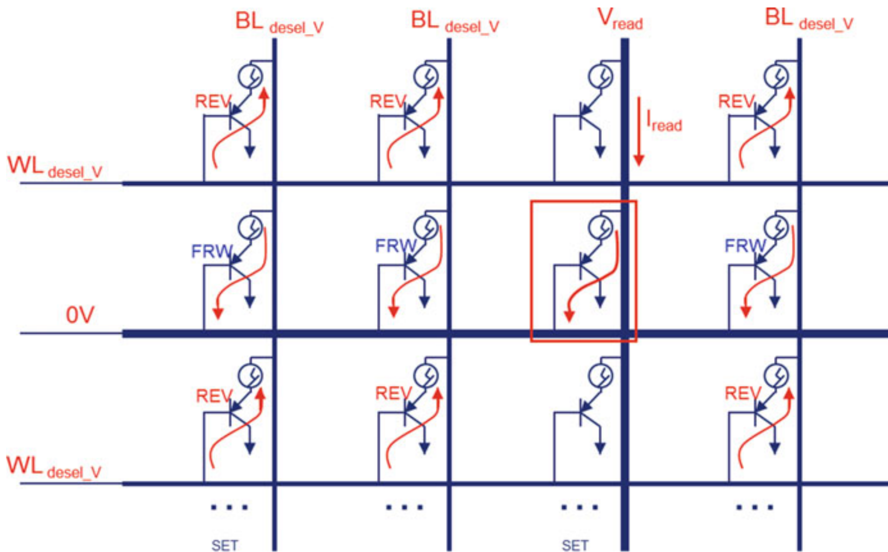


Fig. 10.2 Biasing conditions for selected and deselected cells

In order to be ready for a read operation, each idle tile is biased so that all its diode selectors are reversed biased. One possible implementation is to have all of the word lines biased at the read voltage ($WL_{desel_V} = V_{read}$) and all the bit lines left to ground or floating. When the cell is selected, one WL will be driven low and one or more BLs will be driven high, so that the GST structure will see the voltage applied to the BL minus the voltage drop on the diode selector.

10.1.1 Tile Sizing

The most critical item in the definition of a PCM array architecture is undoubtedly the tile sizing. The tile is defined as the smallest array unit comprising of PCM cells and base contact cells only. Each tile requires its own decoding structures, an n-well ring, dummy elements at the array borders, and possibly a local amplifier to enhance the cell signal before sending it to a global sense amplifier. It is therefore obvious that from a die size perspective, it is strongly desirable to use the largest possible tiles since the decoders, dummy lines, and array rings constitute a fixed offset that negatively impact the array efficiency (defined as the ratio of the cell array area to the full die area).

What elements are limiting the maximum tile size?

IR drops on the BL and WL are the first item to be considered. As the bit lines and word lines grow longer, the IR voltage drop on the lines gets larger, determining:

- An increase of the program voltage that has to be delivered to the array in order to reach the desired program current, with a corresponding increase in power
- The need for a higher deselection voltage on the word lines, causing in turn a strong leakage increase
- A strong unbalance of the read/program voltages for close vs. far cells with respect to their line drivers

Another important aspect is the larger capacitive coupling between lines that we get when the tile size grows, which slow down the read latency, but usually the previous limitations are strong enough to determine the final tile sizing.

So how do we decide what is the maximum WL and BL lengths that minimize the die size without impacting too much the performance? Of course the specification plays a key role here, depending on the focus being on power or speed or on cost, but as a rule of thumb, a certain leakage to program power ratio can be used as a criteria for sizing the tile. For example, let's assume that the power allocated for program is split in this way: 90% for the actual program operation and 10% to compensate the unselected diode leakage. Knowing the reverse leakage as a function of the reverse bias voltage, the sheet resistance of BL and WL, and the program voltage across the selected cell, we can easily determine the largest possible tile size that complies with the above requirements.

It should also be considered that the best array efficiency could be obtained with a non-square shape for the tile. For example, if the sheet resistance of BL and WL are significantly different, a shape with longer lines where the sheet resistance is lower is preferable.

10.1.2 Partitions

After the tile definition, a second key architectural aspect of a PCM array is the use of partitions. Where does the need of partitions arise from?

In read mode, the WL deselection voltage must be high enough to prevent a forward biasing of the deselected diodes on the selected BL. A simple solution would be to bias deselected WL at a voltage equal to the BL read voltage.

In program mode, the minimum WL deselect voltage must be higher, because the BL voltage is now forced to that of a program operation. Theoretically, the program biasing conditions are compatible with the read mode requirements, meaning that no forward biasing of unselected diode will occur; however, the use of such a high deselection voltage has several negative effects:

- The overall deselected cell leakage grows substantially, causing power issues.
- The leakage of the deselected cells on the same BL can become so high that a significant change in the sensing operation is caused and possibly a read error can happen.
- Unless the deselected BLs are forced to ground, the leakage of the deselected cells is collected on the selected WL, potentially causing an IR raise issue. WL raise will further negatively affect the read precision.

For these reasons, it is preferable to use two distinct biasing conditions for read and write; as a consequence, there is a relatively large amount of power that needs to be spent in order to prepare a tile for the programming operation, corresponding to charging all the WL to the higher voltage (Table 10.1).

The need to change the biasing conditions of the array from read to write is determining the other key architectural aspect of a PCM array. In order to save energy, and to switch as fast as possible between the two modes, it is desirable to segment the array in “partitions” as pointed out before. A partition is the portion of the array that will change its bias when prepared for the write operation, while all the remaining cells will stay in read bias.

A partition must contain a sufficient number of tiles to ensure the desired write bandwidth. The reason for this is that each tile can sustain only a certain number of bit programmed in parallel, and usually this number is lower than the read parallelism. The programming current is substantially higher than the read current, and if the WL resistance is high, a large number of bits programmed in parallel would determine a large IR raise in WL voltage. This in turn would cause a raise in the BL program voltage, with a further increase in the leakage figures. Knowing the desired program bandwidth, the average program time and the maximum number of bits per tile will determine the minimum partition size.

As for the tile, partitions too have their own dedicated peripheral circuits, like switches, voltage regulators, as well as fixed borders between adjacent partitions. For this reason, for die size reduction purpose, it’s desirable to use the largest possible partition size, but this is in contrast with power and speed consideration:

Table 10.1 Example of biasing voltages for BL and WL during idle, read, and program

	Idle	Read	Program
Selected BL	–	1.3 V	5 V
Deselected BL	0.6 V	0.6 V	0.8 V
Selected WL	–	0 V	0 V
Deselected WL	1.3 V	1.3 V	4.5 V

the larger the partition, the larger the amount of energy spent for programming and the longer the time to enter and exit the program biasing condition.

10.1.3 Global Word Line Decoder

Decoding a single word line in a tile requires a certain number of signals. Typically the address bits that identify a WL in a tile are converted into pre-decoded signals that can be used directly by the tile row decoder; signals might be level shifted to a higher voltage when program operations are in progress. In order to simplify the routing over the array and reduce the number of circuits needed for the decoding operation, some global decoder signals might be used by the tile WL decoders. Global signals are generated within the partition-specific periphery and are distributed at the proper level shifting as required by the operation in progress.

Figure 10.3 shows an example of this concept. Several global WL run horizontally across several tiles, and each row decoder is simplified because part of the decoding is deferred to the global partition decoder. Similarly the column decoder can be simpler by using many global BL running vertically and one global decoder that is shared across many tiles. This organization will likely require the availability of four metal layers for the array, two for local/global WL and the other two for local/global BL.

10.1.4 Special Program Modes

One interesting note concerns the comparison of the energy spent for the pure cell programming and the energy used by the partition biasing switch. Especially for large partitions and small granularity program operations, the balance can be rather disappointing, with the largest part of the overall energy devoted to the biasing switch. A possible work-around to this problem is a new program mode in which the system SW, knowing in advance that a stream of program operations is going to happen, forces the device to set a partition in the program bias conditions until the stream is completed, getting a remarkable gain in energy and program latency.

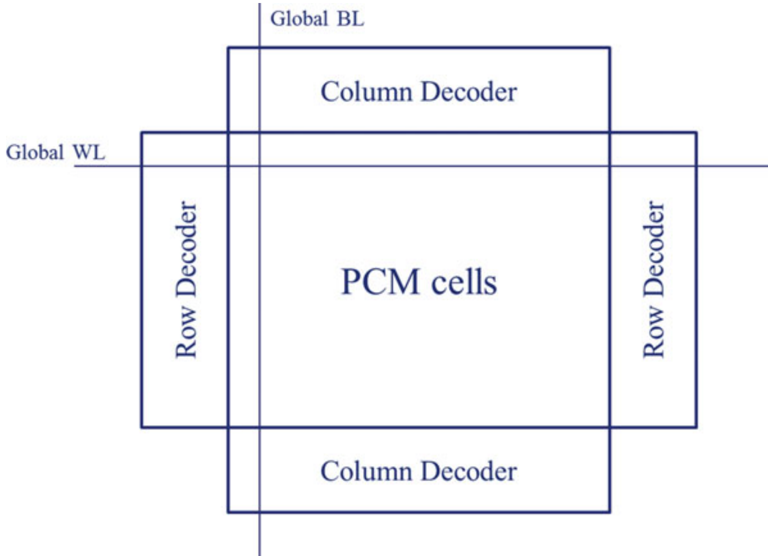


Fig. 10.3 Tile with global word line and bit line signal crossing

10.1.5 Sense Amp Hierarchy

As said above, each tile is completed by its own decoding structure. A large group of columns are mixed into a single line by the column decoding. This line can be directly connected to a sense amp, but in this case, the array efficiency will be severely impacted since the sense amplifier footprint is not negligible with respect to the overall tile size. A much more effective architecture uses global lines running across two or more tiles, with shared sense amplifiers. This scheme allows a compact layout with a limited penalty on speed: the global line capacitance is small (line pitch is relaxed), and few transistors are loading the line, typically one access transistor per each tile sharing the global line.

Program load circuits can also be shared across multiple tiles in a very similar way.

10.1.6 Read While Write/Banking

For large arrays, it is not convenient to extend the sense amplifier or program load circuits across the entire die. Even using relaxed pitch lines, the lines would become long enough to adversely affect read latencies, and there could even be a concern of IR drops on the global lines. So there is a certain convenience, especially if the device spec is performance sensitive, to create several sense amplifier banks and keep the global lines short. In this scenario, a new feature is potentially available at

a relatively low cost to enrich the device specification: it could be possible to manage simultaneously a program operation and a read operation. The big advantage of this approach is that the device is not totally locked out to read when a program operation is in progress. Especially for buffered programs (see Program Operations paragraph), the latency to carry out the program operation might be of several hundreds of microseconds, and this is clearly a limitation compared to a DRAM device, for example, where read and write latencies are comparable.

Since we have multiple sense amplifier banks, it is possible to use one bank to access a partition biased for read, while at the same time, we have a different partition set for program; the only limitation to the read accesses would be relevant to the partition with a program operation in progress; that area would result as not accessible to read.

If a sense bank is serving multiple partitions, it is still possible to manage a read-while-write operation, at the condition that we use separate global lines, one for reading that is connected to sense amplifiers and the other for writing, connected to the program loads.

Can the presence of multiple sense banks lead to offer independent banks in a similar way as DRAM do? In principle this is possible, but this is strongly dependent on the approach used to manage the program operation. In case a program-verify loop approach is followed, and the complexity of managing this operation is assigned to an embedded microcontroller, full bank independence cannot be guaranteed, unless an instance of the embedded microcontroller is available to each bank.

10.1.7 Redundancy

As with any other large memory array, PCM devices must rely on redundancy to achieve good yield/die cost.

The most obvious defects needing redundancy are column and row shorts within a tile.

Row redundancy is implemented by adding a few spare rows (word lines) in each tile. Those rows will be accessed using dedicated row redundancy signals, while the normally selected row is disabled by the redundancy activation.

From the moment the address is delivered from the I/F (interface) to the core, it will take a few ns to detect if a row needs to be repaired, so there is a certain delay associated with row redundancy. However, the activation of a word line just involves its being tied to ground and is usually faster than the column selection, where a precise voltage settling is required; for this reason, even considering the address matching delay, the row redundancy does not impact the read latency.

Since each tile contains a set of spare rows and the number of tiles in a large device is high, the number of potentially replaceable rows is pretty high, probably beyond the minimum required by yield considerations. Having the full flexibility on row repair has a cost:

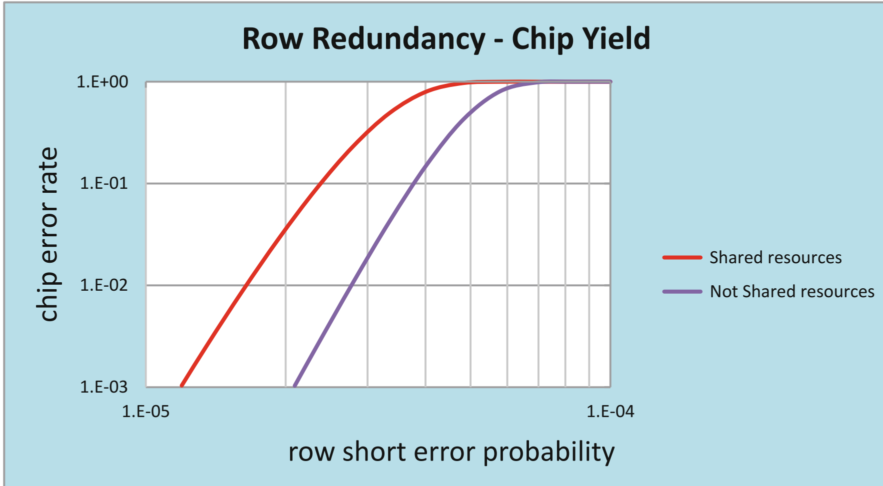


Fig. 10.4 Chip yield as a function of row defect density. The two curves refer to different approaches for defective address storage

- Number of fuses to store the defective addresses
- Size of the address matching logic and routing of the redundancy control signals

For this reason, it is preferable to opt for a compromise in which several tiles will be repaired simultaneously, sharing the control signals and reducing substantially fuses and matching logic (Fig. 10.4).

For column redundancy, there are more options available. A possibility is to use an approach similar to row redundancy, in which spare columns are added to each tile. A more interesting solution consists in adding a complete set of spare columns and dedicated sense amplifier(s) residing in separate tiles. The advantage of this solution is that the address matching delay can be completely hidden behind the sensing phase. Instead of choosing a spare element when the address matching is true, a subset of the spare columns are read concurrently with the standard array columns; during the sense evaluation, a redundancy mux will get ready to replace one or more sense outputs with those coming from the spare sense amps, so the only additional delay is that (very small) coming from the muxing operation.

A third, less obvious defect type is a column-row short. For example, a WL strap contact could be shorted to a BL due to a defect. In principle this defect can be repaired by replacing both the column and the row shorted together, but there is a point that must be taken into consideration.

Even if the two faulty elements are repaired, they still play a role when other columns and rows are selected within the same tile (Fig. 10.5). Let's suppose that a tile is biased for a program operation: the faulty row will be biased at the WL deselect voltage like any other deselected row; this bias of course will propagate to the shorted column. If the deselected columns are biased to ground (or to a BL deselect voltage), current will flow through the BL to ground or to the deselect

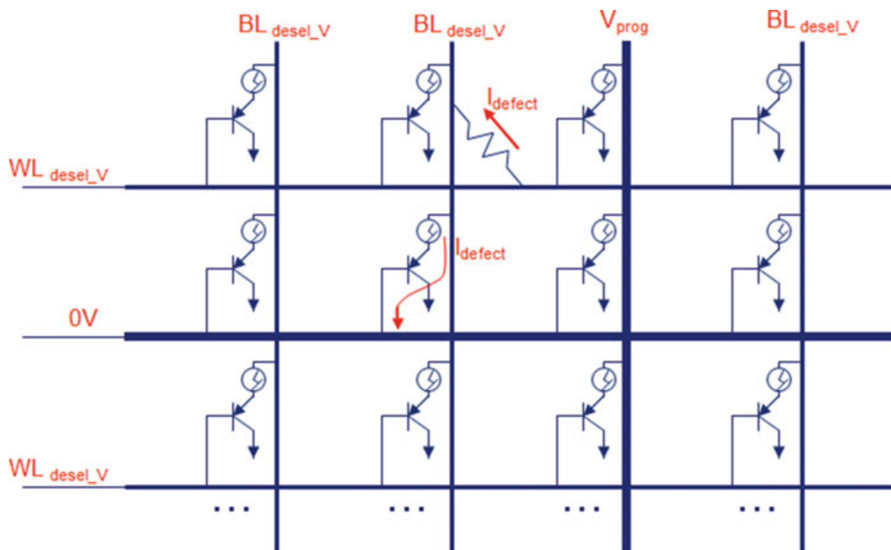


Fig. 10.5 Column-row short, causing sneak current paths even when non-repaired addresses are selected

voltage; if they are left floating, then the shorted BL voltage will raise until current will flow through the cell into the selected word line. In a nutshell, whenever a WL-BL short is present in a tile, regardless of its being repaired, all program operations will see an extra-current load through the shorted elements. Read is likely not affected because of the lower voltages used and the fact that the voltage propagation happens through two diodes. Especially if the short involves more than one line, for example, because the defect is large, it is recommended not to use that tile. So a third type of redundancy can be considered, the tile redundancy, where an entire tile is replacing one from the standard array.

Tile redundancy is probably the most expensive repair in terms of die size, and this is not just because of the obvious consideration that a tile is much larger than a group of columns or rows; there is also an implication connected with the tile biasing. As mentioned above, partitions are provided in order to limit the amount of energy needed to switch a certain array section from read to program bias. A first solution consists in adding a redundant tile to each partition; the biasing of the spare tile is relatively simple, being managed by the same circuitry that drives the standard tile, but the cost of adding one tile per partition is high, e.g. about 6% for a 16-tile partition. Besides, the flexibility of this solution is low, because there can be many unused tiles, but still repairing one partition with two faulty tiles is not possible. Conversely, another solution is to add totally independent tiles that can be associated to any partition, but this requires a complex management of their biasing; dedicated bias circuitry and sense amps for the spare elements must be added, and they might need to be specialized as they drive a totally different load, e.g., a single tile rather than an entire partition (Figs. 10.6 and 10.7).

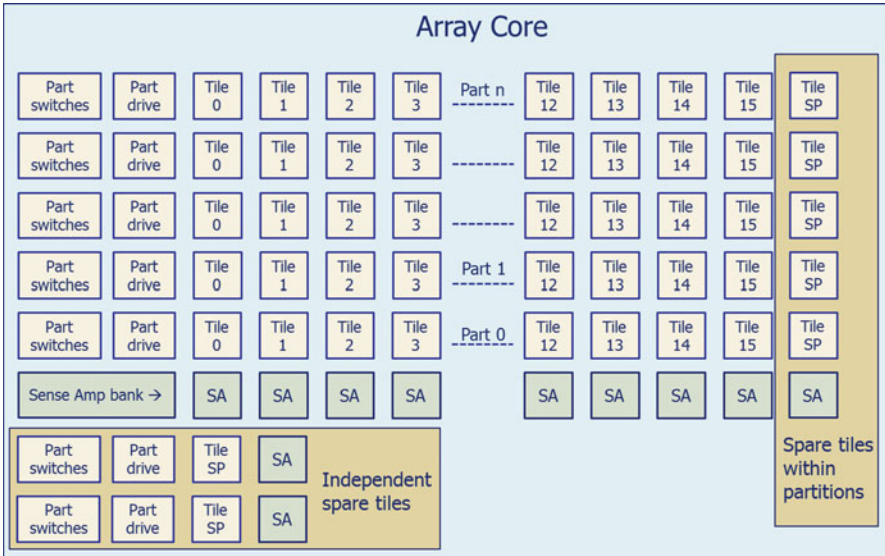


Fig. 10.6 Two solutions for spare tiles: rigidly associated to a partition (*on the right*) or fully independent (*at the bottom*)

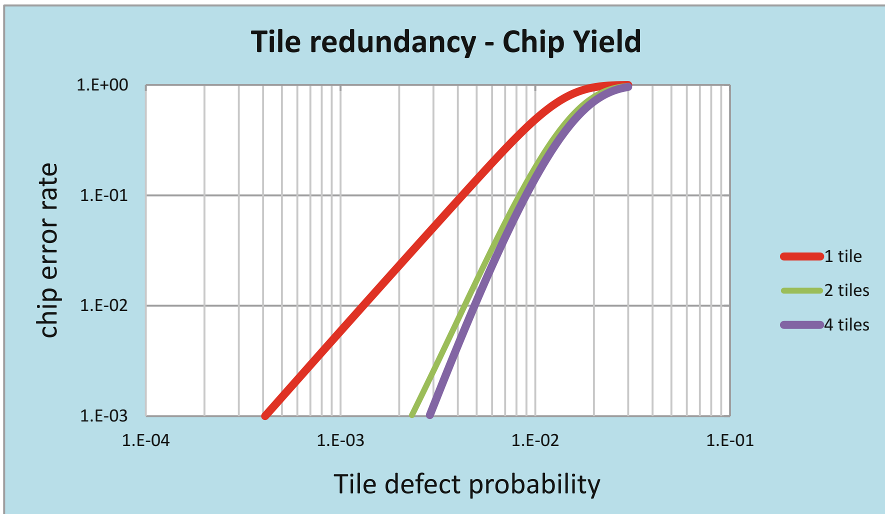


Fig. 10.7 Chip yield as a function of the number of defective tiles and spare elements available

10.2 Array Biasing

As said above, for a proper array functionality, all the diodes in a tile will have to be reversed biased, with the exception of the selected cell(s) whose diodes will be forward biased. Taking as a reference a square tile of n rows and n columns where we address a single cell:

- Roughly n^2 diodes will be in reverse bias conditions: word line voltage high and bit line voltage low.
- $2*n$ diodes will have little or no bias: both diode terminals at low voltage on the selected WL and high voltage on the selected BL.
- One diode will be in a strong forward bias condition: word line low and bit line high.

One possible implementation to achieve this is to have all of the word lines biased at the read voltage and all the bit lines left to ground or floating. When a cell is selected, one WL will be driven low and one or more BLs will be driven high, so that the GST structure will see the voltage applied to the BL minus the voltage drop on the diode selector.

The deselected BL can be either grounded or slightly higher than ground. The use of a nonzero voltage on the deselected BL allows a reduction of the reverse leakage.

In fact, as the voltage on the deselected BLs is raised, a corresponding reduction on the reverse bias voltage is seen on $\sim n^2$ cells, at the cost of an increase of the forward bias voltage on n–1 cells belonging to the selected WL (Fig. 10.8).

Following a very similar line of thought, it is not strictly necessary to use a WL deselection voltage equal to the BL read voltage. This voltage ensures that there is no voltage applied to the deselected cells belonging to the same BL; it is likely that a slight WL voltage reduction is possible before the deselected cells would drive a forward current that is high enough to disturb the read.

10.2.1 Tile Biasing

A simple way of biasing the tile is to use a full P/N decoder to drive both the deselect and the select voltages to the array. Of course this implies a more complex decoding scheme that is affecting the die size, especially if a double-sided decoder approach is followed. There are other smart solutions to achieve the same functionality in a very compact layout.

One is to use a pull-up resistor per WL, with the other end of the resistor connected to a common node that is driven to the deselect voltage. The resistor is easily integrated into the array itself, as a part of the word line active area. The advantage of this solution is that the selection can happen by means of a simple

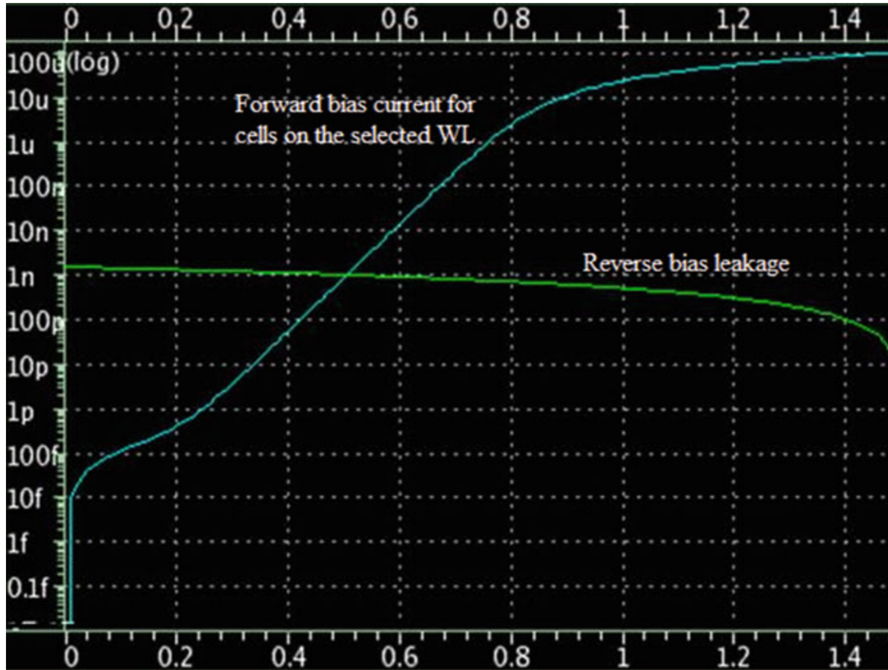


Fig. 10.8 As the deselect BL voltage is increased, the overall reverse bias leakage is reduced. Considering one deselected BL, the intercept between the two curves is the point where the n–1 diodes attached to a deselected WL draw the same current as the diode on the selected WL

NMOS pull-down. When all pull-down are switched off, all the word lines go to the deselect voltage and the tile is in the idle state (Fig. 10.9).

A second solution is offering a compact layout and a way to drive the ideal deselect voltage on word lines.

Figure 10.10 is showing a tile in which a special BL "DBL" is biased at a WL_desel voltage. The diodes connected to this special BL will drive a forward current into all the WL in the tile, until a voltage high enough is reached on the deselected word lines. If an array BL selected for reading is then biased at the same voltage as the DBL line, the amount of forward current flowing into the deselected word lines will be similar to the one we have on DBL. In order to guarantee some margin, DBL might be biased to a voltage that is slightly higher than the read voltage. The advantage of this approach is that the WL deselect voltage will be exactly the minimum necessary to reduce the forward leakage to an acceptable level, therefore minimizing the reverse leakage on the deselected cells. Following a similar approach, all unused bit lines can be driven by a special word line DWL which is attached to deselected columns (e.g., a WL connected to odd cells when an even address is selected). The unused BL will be drawn close to ground at exactly the value that reduces the forward conduction on the selected WL to an acceptable level, therefore reducing further the overall reverse leakage.

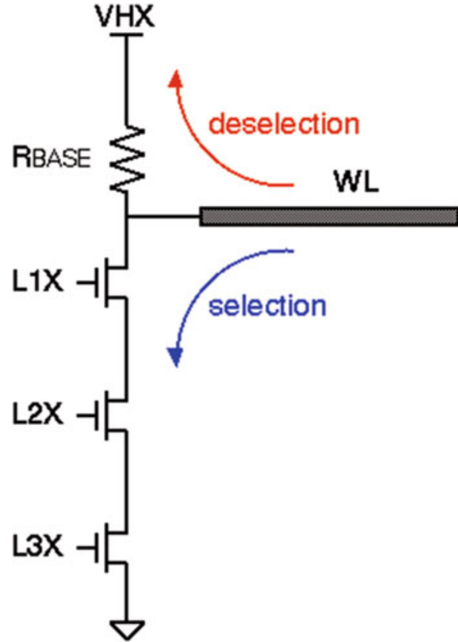


Fig. 10.9 The simplest row decoding approach, with N-channel pull-down MOS and a pull-up resistor integrated in the array

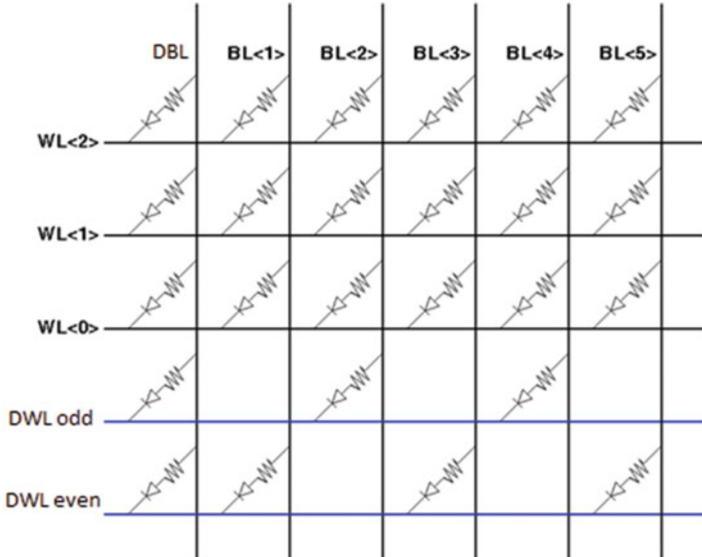


Fig. 10.10 Dummy line biasing. A DBL column biases all unselected word lines, while two DWL rows bias all unselected odd or even bit lines

It must be considered that both solutions, the pull-up resistor and the special BL, will inject current into the selected WL, causing an extra power consumption and a voltage drop on the pull-down device of the decoder. It would be best to have the pull-up resistor or the special BL on the same side of the pull-down device, or there will also be an additional IR raise across the WL.

10.2.2 Line Coupling

There is one concern referred to fine pitch arrays relative to adjacent line capacitive coupling. A strong line to line coupling is not easily managed by the smart decoder solutions described above. It is essential that, if one line (row or column) has been selected, the other does not affect the adjacent lines while switching, or a spurious write might occur (at least in program mode). For this reason, at least one of the two directions should be managed by a full CMOS decoder, unless slow switching speeds are acceptable.

10.2.3 Voltage Compensation

As stated above, the use of long lines determines an unbalance in the effective voltage applied to the cells. Given a certain BL/WL resistance and a program current, there will be a voltage drop on the selected BL that goes from nearly zero when a cell close to the BL driver is selected, up to I_program*BL_res for the furthest cell on the opposite side of the tile. This is an undesirable situation, because in order to guarantee an effective program voltage for the far side cells, we risk to overprogram the near side cells.

One possible mitigation factor for this problem is to design a double-sided WL/BL decoder, so that the program current is delivered through both line ends, thus dividing by four the voltage drop. Of course this approach has an inherent die size cost.

Another effective method to avoid overstressing the near side cells is to apply a voltage compensation factor, following the array topology. If the voltage drop is Vd, we can reduce substantially the overvoltage by splitting the tile in regions which will receive an adjusted voltage depending on how far each cell is from its driver. Only the far cells will see the highest voltage, while the near cells will be biased with a scaled down voltage. It must be noticed that this methodology is easily applied to the BL side, but not as easily to the WL side if more bits are programmed simultaneously; each bit has a different distance from the WL driver so it’s not possible to compensate the different IR drops simultaneously.

Another point worth noticing is that a full compensation is not possible: the real amount of IR drop depends on the actual program current, which has its own dispersion, so the voltage on the furthest cells will be only partially compensated.

10.2.4 Voltage Biasing vs. Current Biasing

The above considerations raise a question whether a different approach in biasing the PCM array is more appropriate. Rather than using a voltage biasing, we could decide to force a current into the selected cell. This choice is very effective in neutralizing the IR drop variations between close and far cells, because the same amount of current is injected in all cells regardless of their topology in the tile. The drawback of this approach is that the geometrical variations of the cells are not taken into account. A large-diameter cell will be underbiased, while a narrow-sized one will likely be overstressed. The voltage bias approach instead is friendlier toward geometrical variations, because the current flowing into the cell tends to adjust to the different geometries. As a conclusion, for large tiles without voltage compensation, the current biasing can be an effective way to control the program biasing, but if a voltage compensation is applied and/or intra-die geometrical variations are high, then the voltage bias approach should be advantageous.

10.2.5 Standby Biasing

A special care must be paid to tiles in idle mode. The idle state can be maintained for a long period of time, during which the diode reverse leakage can slowly charge the bit lines to a voltage that is progressively approaching the word line voltage. This slow drift of the bit line voltage has one positive implication: diode leakage goes down as bit line voltage goes up. But at the first tile access after a long standby, with the selected word line going low, there will be a very large capacitor, constituted by all the bit lines of the tile, that has to be discharged through the word line itself and its own decoding path. The injected current might be very high, causing a high IR raise and making the read impossible until the discharge process has been completed. Since most specs require the same read access time regardless of how long is the interval occurring between two read cycles, it is necessary to provide a bit line voltage clamp that prevents this phenomenon. The "dummy word line" scheme described above is one way to effectively provide this clamp and in particular to provide a clamp that discharges bit lines at the highest possible value, minimizing leakage and therefore standby current at any device temperature.

10.2.6 Power/Energy Considerations

As mentioned above, the tile size will affect dramatically the power figures of a PCM device.

First of all, let's consider that the high voltages required in program will likely be generated by an internal charge pump. Charge pumps are not very efficient (they usually have a power efficiency well below 50%), and their efficiency tends to

worsen as the output voltage grows higher and the number of pumping stages goes up.

Considering an ideal cell with a set current of 100 μA and a voltage across the diode + GST series of 2 V, the power sunk in program is 200 μW. Reality is much worse though.

First of all, let's add all the IR drop components. BL and WL resistance could be in the 5 KΩ range each; WL and BL decoder MOS resistance might be of that order too; additional 5 KΩ could come from contact resistance (emitter resistance, WL strap, or base contact resistance); 2 KΩ could be associated with the global BL resistance and the program load. In practical terms, the worst case cell will see ~27 KΩ series resistance corresponding to a 2.7 V IR drop on lines and decoders. Besides, if 100 μA is the average set current, the voltage will have to be raised further to guarantee that a large cell still has enough biasing; assuming a sigma of 10 μA for the set current, the device must be ready to sustain an IR drop of 4 V. The desired output of the charge pump is therefore set to 6 V (4 V + the voltage across the cell). Assuming a power efficiency of the charge pump at 35% with this high output, the real power absorbed by this average cell is therefore 100 μA * 6 V/ 0.35 = 1.7 mW, i.e., almost 10× the initial estimate. And this calculation does not take into account the energy wasted due to array reverse bias leakage, the tile bias switching, the embedded firmware execution, and the possibility that multiple program/verify loops are executed. Using smaller tiles will at least reduce the IR drops allowing a lower program voltage and a better charge pump efficiency.

10.3 Sense Amplifier

To achieve a good read signal, the first thing we need to determine is the best read bias voltage for the BL. The cells consist of two series elements, a diode and a GST memory element; the voltage we drive on the selected BL (assuming the selected WL is grounded) goes across the diode-GST series, but there is no way to predetermine the voltage split between the two, as this depends on the GST status. It is important to notice that the diode has an adverse effect on the cell bias, since the GST memory element will see a voltage that is lower on the set, low resistance state cells, therefore reducing the read signal.

The maximum voltage applicable across the GST structure is limited by reliability concerns, but we should try to get as close as possible to a safe value (V_safe) in order to extract the maximum signal from the cell. In fact, due to the exponential characteristic of the diode, the voltage across the diode itself does not change much for a given increase of the biasing voltage, so that the extra voltage is delivered almost entirely to the GST element.

How is the read voltage regulated to achieve the maximum allowed bias? The proposed approach is to measure the voltage drop on a reference diode and then add the V_safe voltage to generate the BL bias voltage.

A dedicated mini-array containing the reference diodes is biased with an "I_safe" current, defined as the maximum current allowed through a reset cell

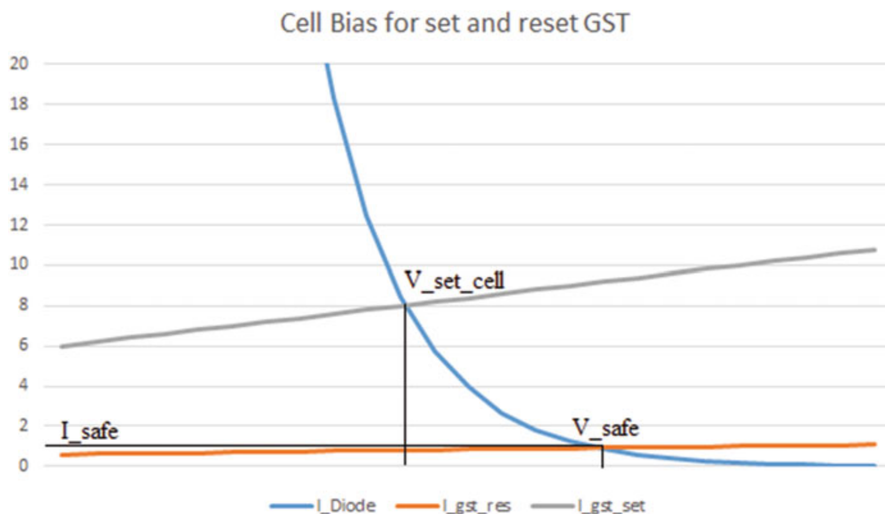


Fig. 10.11 For the set cell, the voltage across the GST is lower than V_safe, by the amount required by the diode to drive the additional current with respect to I_safe

before incurring reliability issues. Several diodes could be used in parallel to get a more stable measure. A voltage adder is then adding the reference diode(s) voltage with the V_safe voltage. Reset cells will see the full V_safe across the GST but no more than I_safe current, while set cells will see V_safe minus the additional voltage drop on the diode requested to drive the cell, as a result by the intercept of the diode characteristic with the GST load line (Fig. 10.11).

This represents the biasing voltage that guarantees the best signal without incurring read disturb issues.

V_safe itself can be generated in several ways. The maximum voltage tolerated by a cell depends on temperature and process variations, so it must be provided a mean for a die-by-die trim of this value, on top of a temperature-dependent regulation. One simple approach to achieve both regulations is to use a look-up table for V_safe, with an on-chip temperature sensor and V_safe values that are decided during testing. The value corresponding to the current temperature would then be driven into a DAC to generate the desired V_safe value.

Once we have a good biasing voltage for the BLs, the following step is to measure the current flowing in the BL under these biasing conditions. A second mini-array containing several cells in the set state can be used to generate the reference current. Several cells are measured in parallel to reduce variations; they are under the same biasing conditions as the array cells, and they are taken with a ratio that is selectable at testing in order to maximize the read signal. Such current is then distributed across the die and sent to all sense amplifiers. Theoretically reset cells could be included in the mini-array too, to achieve a sort of self-adjusting optimum reference value between set and reset currents; the problem with this approach is that the reset drift may change radically the reset current over time, so it is not easy to exploit a reference cell that was put in the reset state at testing or even refreshed periodically.

The cell current and the reference current are typically converted in voltages within the sense amplifiers, and following that, a simple comparator will determine the cell state. A latch will usually follow the sense to retain the digitalized result of the read operation even after the BL has been reset to the idle voltage.

An interesting consideration about the sense amplifier is the following: for a given tile, two or more active bit lines can insist on the same active WL. This will determine a pattern-dependent read window loss, due to the current of the set cells flowing into the WL. Let's consider, for example, a situation in which all the cells read simultaneously in a tile are in the set state; several cells have a very low resistance state, while one is only marginally set. The best set cells will drive a remarkably high current in the WL, thanks to their low resistance value, causing a certain voltage raise on the WL itself. This does not represent an issue for the good cells but is creating a problem to the marginally set one: on top of having a resistance which is in the high end of the set distribution, its bias is marginal due to the WL IR voltage raise. For this reason, an effective sense implementation is one which causes the immediate BL reset after a "set" state has been determined. The high set current flow is immediately cut off, and the corresponding WL voltage returns to a lower value, easing the read of marginally set cells. The only drawback of this implementation is the longer sensing time needed to let the WL voltage settle to the final value (Fig. 10.12). Whatever is the read scheme, there will always be a

Fig. 10.12 A weakly set cell, initially not capable of triggering the sense amp, is read correctly after the other cells on the same WL have been cut off

pattern sensitivity if more bits are read in the same tile, and this must be taken into account at testing.

10.4 Program Operation

The advantages of the PCM cells over flash cells are the write speed, especially for the reset operation, and the bit alterability that makes the need for organizing the array in erasable sectors obsolete.

In order to fully exploit the potential of the technology, it is very important to put in place an efficient management of the program operation.

Two well-distinct methodologies can be followed:

- A classic program-verify loop approach, similar to those used in flash memories, which can be fairly complex but allows an optimization of the RBER and of the write energy per bit
- A single-shot approach, aimed at getting the fastest write performance but that requires most likely some form of media management

Let’s consider the program-verify approach first.

First of all, it is almost unavoidable to refer to an embedded microcontroller to manage the complex sequence. The use of a state machine, for example, would not have the same flexibility as the microcontroller posing serious obstacles to the device validation/engineering.

The microcontroller is programmed to execute an embedded firmware that has access to all the analog blocks of the device, as well as the addressing of the array. If a “read-while-write” architecture is used, the microcontroller will drive the write addresses, while the read address will be communicated by the user through the device I/F.

A typical program flowchart would be the following (Fig. 10.13).

A certain number of words to be written are loaded through the external interface into a program buffer. The use of a program buffer saves us from the need to repeatedly switching on and off program conditions and allows instead a larger chunk of data to be written at once.

When the program buffer load is complete, a first verify loop is performed on the addressed cells; a second “modify” buffer holds the results of the verify operation, with a “1” indicating that cells are already in the desired state, so they don’t need to be programmed, and “0” showing the opposite. The first verify is therefore saving a 50% cell writes in case of a random over random pattern write.

The first program loop is then started; in order to save power and time and enhance reliability, the shortest pulses with the lowest program voltages are initially applied to the cells. A second verify loop is then performed, clearing the bits in the modify buffer as the verify operation reports the desired cell state. Loops are repeated increasing the pulse duration and/or amplitude until all bits are set in the desired state or the maximum number of loops has been reached. Once a bit is verified correctly, it is removed from the “modify buffer” list, so that the possibility

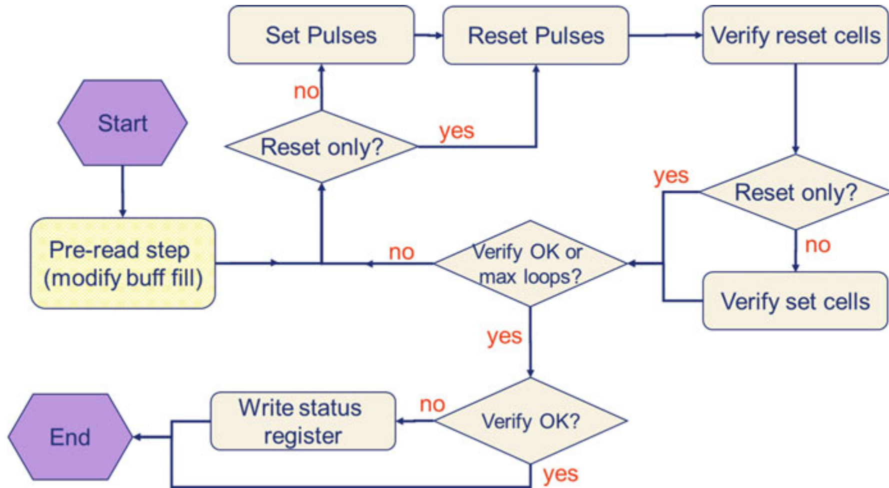


Fig. 10.13 A simple program flowchart showing the main operation blocks

of repulsing it if a subsequent verify fails is excluded. Repaired columns are also cleared from the modify buffer and replaced with the corresponding spare elements; this step is very useful in preventing a large current surge when, for example, trying to program a column that is shorted to ground.

One important remark is that in order to save time and have each loop executed faster, in spite of the longer program pulses, a HW program accelerator is capable of skipping any group of bits that do not contain any cell to be programmed.

Each verify is done using margins, i.e., the reference currents used in verify are more severe than those used in a normal read, thus giving a certain confidence that the following read operations will pass without errors.

Since the set to reset transition is usually much faster than the opposite one, there is a certain advantage in writing a pattern over a preexisting "all 1" background ("1" is by our conventions the set state coding).

A dedicated algorithm can be developed to treat this special case and get the fastest program performance.

For the normal case where both "1s" and "0s" are to be programmed, it is worth noticing that the calibration of the pulse duration is very important in order to achieve the fastest program time. In fact, while the cell reset is not very sensitive to the pulse duration, the set operation is. Choosing too short a pulse in the first loop will minimize its duration but will carry a larger number of bits needing the second, longer pulse. If the pulse is uselessly long, only a small fraction of cells will undergo the second pulse, thus minimizing its duration, but the first loop will become dominant in the overall program latency calculation. Therefore, it is very important to characterize the behavior and even choose a die-by-die trim to optimize the parameters. A "program calculator" tool can also be developed to simulate the device behavior as a function of its configuration and specific cell parameter distribution.

The other program management approach is based on using fixed amplitude and duration pulses, without any verification before or after the pulse. The first consideration is that this approach is much simpler, does not need sophisticated sequences, and is also faster in the sense that it covers everything with a single event (rather than looping multiple times); the verify operation itself takes time and energy to execute.

The very big disadvantage of the "blind program" is connected to the widespread of the set operation latency. If the difference between the minimum set pulse needed by a fast cell and a slow one were small, then opting for the longest pulse on all the cells with the blind program approach might be advantageous, due to the time savings relative to the multiple loops. But if the dispersion of the set pulse is very high, e.g., a 10× ratio or more between short and long pulses, then using the longest pulse for all cells is unacceptable. The work-around then is to trade off the pulse duration for the RBER (Raw Bit Error Rate) of the program operation; by accepting that a certain number of cells will fail, we can use a much shorter pulse, function of the set time distribution, and reduce both power and program latency.

Of course this methodology must be coupled with a strong enough ECC corrector, to bring UBER (Uncorrectable Bit Error Rate) down to useful values.

10.4.1 Program Disturbs

There are several disturb issues connected with the program operations, for example, disturbs induced by cross talk between adjacent lines (word lines or bit lines) or by defects in the tile that can determine an unwanted raise of a deselected line. One particularly relevant disturb is the thermal disturb, in which the deselected cells might change state as a consequence of the high temperature seen in a neighbor selected cell.

Apart from process construction details aimed at preventing thermal disturbs [4], there are also architectural measures that can be taken to mitigate this issue. The use of a program buffer, for example, allows programming several bits at once in a specific array portion. If the bits relative to a program buffer are mapped to a rectangular group of cells made of physically adjacent bit lines and word lines, then most cells will be thermally interacting between each other at each program operation. Only those cells which are on the boundaries of the rectangular section will have physical neighbors belonging to a different logic program buffer. These cells are the most critical because there is a chance that a neighbor program buffer is reprogrammed for a high number of cycles before they are programmed again: so they have to withstand many thermal disturbs while retaining the data originally written. The proposed mapping of the program buffer in groups of adjacent cells minimizes the number of "boundary cells" and therefore makes it possible the adoption of a refresh algorithm that, once the program operation is completed, verifies the state of the adjacent word lines/bit lines. If any marginal cell is found, a new program cycle on the neighbor cells is started to restore a robust crystalline or amorphous state in the disturbed cells (Fig. 10.14).

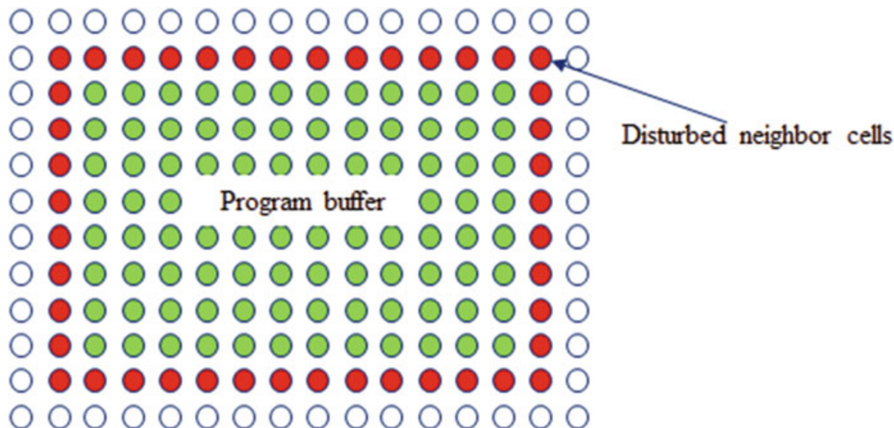


Fig. 10.14 A program buffer addresses a rectangular group of adjacent cells (in green). The red cells adjacent to the program buffers are disturbed by multiple program operations of the green cell; a refresh algorithm can recover the original state of the disturbed cells

10.4.2 Erase Operation

Since PCM cells can be set on a bit-level basis, there is no need for an erase command like in flash memories.

There are still a few advantages though for considering the implementation of an erase command. The first reason is simply for backward compatibility with existing flash devices. Especially in the early stage of PCM device adoption, having a fully compatible specification will ease their introduction into existing systems.

A second reason is to fully exploit the higher speed of the reset operation: in order to prepare a certain memory area for this reset-only write, it might be helpful to use an erase command.

Finally a flash-like management of the array, though having clear disadvantages and not fully exploiting the bit alterability offered by the technology, is also opening the possibility of a sector-based media management, potentially including things like storing the number of cycles in each block or the time elapsed since the last erase, etc.

In all cases, the erase operation will be achieved as a sequence of set pulses, so it is not related to a real bulk operation as happening in flash memories.

10.4.3 Suspend Commands

As explained above, the latency associated with read and write operations is rather different, especially when a set operation is involved. If the device offers an erase

command, its latency can be several orders of magnitude higher than a read. For these reasons, it might be wise to include a suspend command in the specifications. The suspend command will temporarily stop the program operation to allow reads in the same partition; the program is then resumed later until the operation is completed or until a new suspend is issued.

The main actions to be taken to correctly manage a suspend command are the following: first, the partition biasing voltages have to be reset to the read conditions; second, the current state in the program sequence must be saved to allow resuming the operation at a later time. It must be noticed that to ensure that the program operation will come to an end, the interval between two suspend commands must be above a certain threshold or most of the time will be spent to enter/exit the program biasing conditions rather than delivering actual program pulses to the array.

10.5 Multilevel Cells

Since the phase change in the GST material involves a volume that is proportional to the amplitude of the reset pulse, it is possible to envision a program algorithm that tries to step the reset voltage until the desired amount of GST is reset in the amorphous phase, controlling the resistance of the cell in a nearly analog fashion.

The use of an embedded microcontroller in this case is definitely a must, and if the program steps are sufficiently fine, it is possible to get three nonoverlapping resistance distributions, on top of the set distribution, implementing a two bit/cell multilevel cell. One important detail about this flow is that it is absolutely necessary to clear the "modify list" buffer from bits that are passing the program verify. If this is not done, it might happen that one cell has marginally passed the program verify, goes through several program-verify loops, and at one point fails the verify operation. The cell would therefore undergo a further program pulse, but the problem is that after several loops, the program voltage has increased substantially, and a single pulse could now place the cell in a totally wrong distribution.

The biggest problem in the multilevel approach is certainly the drift; the initial placement of the cells may be accurate, but over time they will increase their resistance due to drift. Even assuming that the distributions move consistently maintaining a separation between them (which most likely won't happen because of the geometrical variations in the cells), we still have the problem of using a fixed reference to discriminate a moving distribution. There are work-arounds for this problem, one of them being the use of additional bits that work as a reference in each program page; the sensing operation will adjust its current reference levels until the reference bits are read correctly and use this setting to read the rest of the page. Clearly this is a slow approach (it might require a few read attempts), and it requires a per-page write specification as well as some die size penalty for the reference bits.

10.6 Microcontroller and Embedded Firmware

If the microcontroller approach is used, a proper mean for storing the firmware code must be provided.

One option is using the PCM array itself to store the code, probably in one dedicated array not to interfere with the standard array space. There is a disadvantage in this approach though, as PCM might not be reliable enough to retain the code after the reflow process to solder the device to a PCB board. Besides, we would still need a code to write this dedicated array. A ROM can also be used, solving both issues, but obviously its flexibility is much lower, since a new mask is needed for each code revision. The proposed approach is a compromise between the two: a ROM holds a basic code that implements all the key program routines, with parameters that can be stored in fuses to fine-tune the code behavior. An additional set of fuses is provided to apply small fixes to the base code, without having to store the entire code in the fuses. This decision is important because the area taken by a single fuse is huge if compared to a PCM cell.

On top of the RAM and the fuse code repair mechanism, the device is provided with an embedded RAM that is also usable to execute code. In this way, several variations of the program code can be tried without having to request a new ROM and/or blowing fuses to change the code.

10.7 ECC

The presence of an embedded ECC engine offers great opportunities for improving the device reliability and its yield [2, 3]. Many reliability phenomena (data retention, cycling endurance, drift, to mention a few) exhibit a tail in the parameter distribution that can be easily recovered by the ECC. An example of ECC correction calculations is reported at the end of Chap. 5, based on [2].

Thanks to the relatively high read speed of a PCM device, our preference should go to an ECC scheme capable of doing the correction in few ns, to prevent big losses on the read latency. Schemes based on large code words and high number of correctable errors with iterative ECC engines like those used in NAND flash devices are therefore not recommended. Schemes that operate on 128 ~ 512 bit code words with 1 ~ 3 correctable errors can instead be implemented with a relatively low latency and a small area. In case the device is supporting read-while-write or concurrent reads in different planes, two or more ECC engines must be provided.

The high speed for ECC is required only for the correction phase, since the latency is added to a read operation usually taking less than 100 ns; for program, the latency associated with the tile bias switching and with the set operations easily brings the program duration in the range of several μs, so using a more compact iterative scheme for the ECC parity calculation is perfectly acceptable.

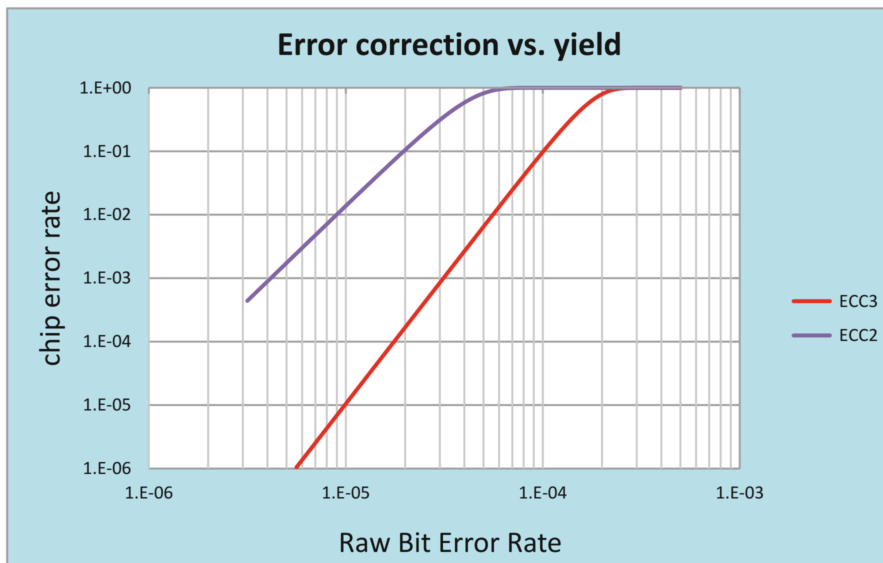


Fig. 10.15 Yield loss for a 1 Gb device at a given bit error rate for BCH2 and BCH3 ECC configurations

ECC also offers the possibility of implementing an embedded refresh command, capable of addressing time-dependent reliability issues like drift and retention in a very effective way.

For example, a refresh command could be implemented in a way that all code words are read sequentially by the microcontroller, and whenever an error correction is detected, the specific code word would be refreshed by the embedded firmware, in an attempt to rewrite correctly the bad bits. The advantage of the methodology is that the array contents are kept refreshed, maintaining the error rate below the ECC correction threshold, writing only the words that need to be updated, and thus saving power and time for the refresh operation (Fig. 10.15).

10.8 Fuses

All memory devices have the need to store dynamically nonvolatile information regarding redundancy, configurations, trims, testing information, and so on. For a nonvolatile memory, the task of storing information is normally not difficult, as they can use directly the memory cells (in a dedicated array) as a storage media. But PCM poses a new challenge, as the data retention is not guaranteed at temperatures as high as 250°C even for a very short period of time; therefore, the soldering process on a PCB board is likely to wipe all information stored so far.

An alternative to storing data directly in a PCM array is the use of a fuse-based storage. Fuses are usually larger than a memory cell, but they are much more insensitive to heat. The number of fuses needed on a device can be of the order of several thousand: for example, it is not unusual to use 15–20 bits to store the address of each spare element used by redundancy; analog trims can involve several blocks, and each of them can take tens or hundreds of bits, etc.

Given the critical role served by fuses, it is very important to ensure a high reliability on fuse data. Methods to improve reliability can vary from voting schemes in which three or more fuses hold one bit to ECC-covered implementations, with fuses organized in pages protected by one or to two error repair schemes. If an ECC scheme is adopted, since fuses are of course not rewritable, it is important to organize data in a way that does not require write access to the same page more than once: in fact it would still be possible to blow unblown fuses in a previously written page, but there would be no guarantee that the new parity could be stored.

Fuses are another example of how the use of an embedded microcontroller can simplify a lot the management of the device functionality. With a proper power-up embedded sequence, fuse contents can be read sequentially, corrected, and delivered to a peripheral block using a common bus controlling when the new data is made active. The sequential download of fuse contents can also prevent large current spikes that would happen at power-up if all fuses had to be read simultaneously (Fig. 10.16).

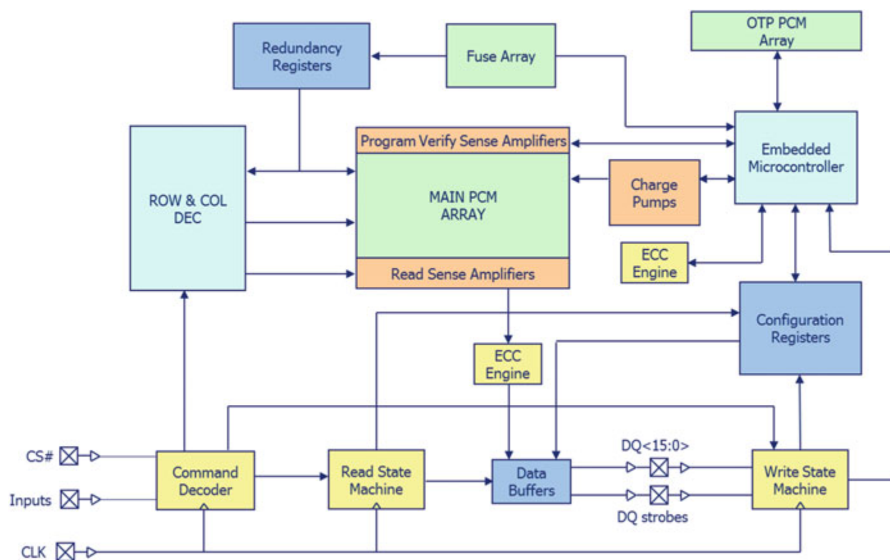


Fig. 10.16 Block diagram of the PCM periphery showing how fuse contents can affect the device configuration by means of the embedded microcontroller; configuration bits are read from the fuse array and deployed to the main periphery blocks

10.9 Conclusions

The key factors influencing a PCM array architecture have been analyzed.

The need for breaking up the array in several sub-arrays, reaching a trade-off between die size, power, and speed, has been discussed, also referencing the effects of features like read-while-write and banking.

An optimum array biasing has been explained, with different solutions privileging specific spec parameters. Sensing and programming operations have been shown, and the benefits of using an embedded firmware were demonstrated.

In summary, PCM is confirmed as a viable technology for a high-density NV memory, probably not as cost-effective as a NAND flash but giving tremendous advantages in terms of endurance and read speed. The bit-level alterability offered by the technology opens fantastic opportunities in simplifying the management of solid-state storage devices and boosts even further the effective write speed of such devices. Power is a critical aspect of the new technology, but there are design/ architecture solutions that can mitigate the problem.

Acknowledgments I wish to thank for their contribution to the PCM core design Andrea Martinelli, Christophe Laurent, Claudio Nava, Daniele Vimercati, Diego Ornaghi, Duane Mills, Efrem Bolandrina, Jerry Barkley, Juri Zambelli, Marco Defendi, Mounia El-Moutaouakil, Pierguido Garofalo, Riccardo Muzzetto, Rich Fackenthal, and Stefan Schippers.

References

1. Corrado Villa, D. M, A 45nm 1Gb 1.8V phase-change memory. ISSCC proceedings, 270–271 (2010)
2. P. Amato, C. Laurent, Ultra fast, two-bit ECC for Emerging Memories. IMW proceedings, (2014)
3. S.B. Paolo Amato, Fast decoding ECC for future memories. IEEE J. Sel. Areas Commun **34**, 2486–2497 (2016)
4. G. Servalli, A 45nm generation phase change memory technology. IEDM Tech. Dig. (2009). https://doi.org/10.1109/IEDM.2009.5424409

Chapter 11
PCM Applications and an Outlook to the Future

Gregory Atwood

11.1 The Importance of Memory in Modern Computing Systems

The performance and functionality of modern systems are being defined as much by the memory as by the CPU. Memory is becoming a significant component of the system cost, a primary bottleneck for system performance, and a major contributor to system power. The impact of memory spans from large systems such as servers down to small systems such as smart watches and even into the emerging IOT (Internet of things) devices.

For high-end systems, such as servers (Fig. 11.1), large amounts of low-latency memory "close," both physically and temporally, to the CPU are critical for the operation of large transactional databases which are becoming prevalent with the growth of the mobile device market and the use of these devices for "real-time" data and financial transactions.

For smaller systems, such as phones, memory is a meaningful contributor to the cost and functionality as well as to the battery life. The importance of memory for this class of devices is illustrated by the fact that, while few people can tell you what CPU is in their phone, almost everyone can tell you how much memory is in their phone as this is often a selling feature. Memory in these devices has been increasing rapidly as their functionality increases to rival that of a laptop computer.

For even smaller devices, such as a smart watch or IOT devices, the memory density is lower, but memory is still of high system importance and plays an important role in defining the functionality and power consumption of the device.

DRAM and NAND memories, the two dominant memory types, have been a primary enabler of the complex systems used today. DRAM, fast but volatile,

G. Atwood (✉)
Micron Technology Inc., Milpitas, CA, USA
e-mail: gatwood@micron.com

© Springer International Publishing AG 2018

A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7_11

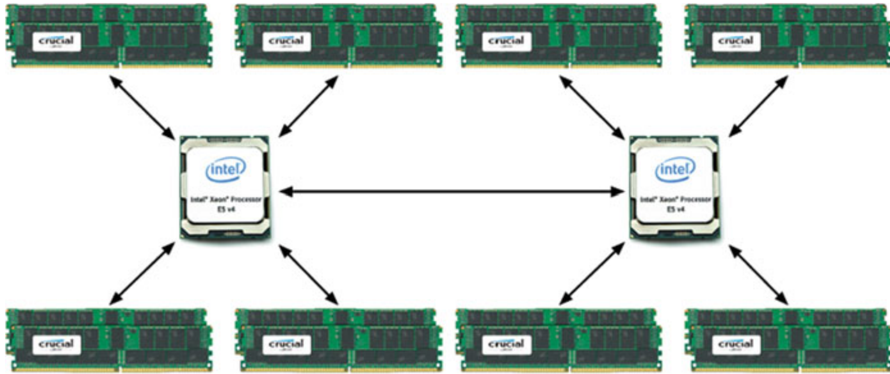


Fig. 11.1 Diagram of a representative (year 2016) 2U enterprise server configuration with 2 Xeon processors and 16 32GB DIMMs of DRAM (512GB). The memory cost is 2.5× the processor cost

serves as the main working memory and NAND, slow but nonvolatile, as the main solid-state storage memory.

11.2 Memory Scaling and System Challenges

Unfortunately, just as the importance of memory is increasing, the ability to continue memory scaling is becoming a concern. Electrostatic storage physics is the basis of NAND and DRAM memories and has been highly successful over the past 40 years. Straightforward scaling of existing materials and structures has enabled a fairly low-risk path to higher memory density, and a pace of roughly two times higher memory density every 2 years has been maintained. Over the past 10 years, though, the complexity of scaling has increased rapidly as we approach some of the fundamental physical limitations for the electrostatic-based storage physics (Fig. 11.2). Significant innovation for materials and structures is now required at each new technology generation. Scaling now requires new expensive tools and an increased risk of major innovations being required at each generation, reducing the financial viability of scaling. With the reducing ability to increase density through smaller geometries, the drive for continued scaling is forcing a transition to three-dimensional structures with new materials and new constraints.

From a system "scaling" point of view, as system performance continues to increase and the amount of memory in the system increases, the memory-imposed system limitations are becoming relevant. DRAM is a volatile device; data is lost on power loss. DRAM requires a frequent refresh of the stored data when power is present; it must be managed by the system. NAND is nonvolatile and serves as a good companion memory to DRAM, but it has a high read and write latency and requires extensive system management to mitigate its high error rate and low

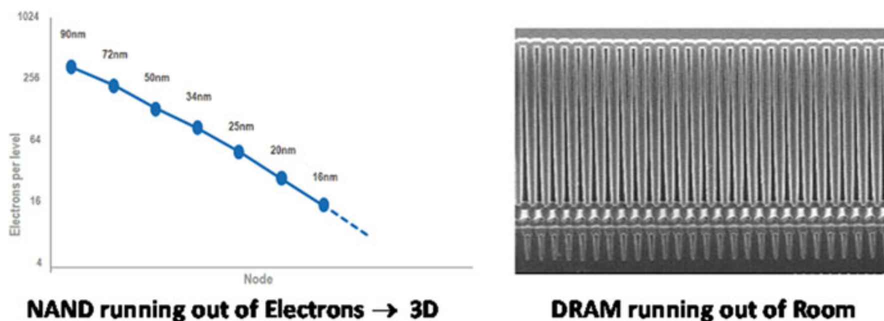


Fig. 11.2 NAND and DRAM scaling challenges [1]

number of allowed reads and writes. Neither DRAM nor NAND are ideally suited to the needs of the system.

The volatility of DRAM requires significant system management to mitigate the risk of power loss resulting in the loss of critical information. Techniques such as journaling and check pointing are used to enable the ability to recover the system in the case of loss of the data stored in DRAM. These operations consume bandwidth (performance) and power by frequently moving data across the system bus to a nonvolatile storage device, often NAND memory. The requirement to refresh the DRAM (rewrite the stored data) in order to preserve data is independent of if the DRAM is being used or not and impacts both power consumption and performance, both issues which increase as the DRAM density grows. Power is consumed by the refresh operation even when the DRAM is not being used. Performance is impacted when a read or write is requested to a DRAM location that is undergoing refresh.

The high read/write latency for NAND and the extensive management that is required to maintain NAND reliability consume both bandwidth and power. Increasingly complex NAND memory controllers can consume as much energy per bit of NAND as consumed by the NAND. Read and write latencies to NAND are 1000 times longer than DRAM and are nondeterministic to the system, depending on the amount of management required.

For additional information related to memory scaling challenges, refer to references [1–4].

11.3 The Memory Hierarchy Today and the Opportunity for Insertion of New Memory Types

The environment of memory physics and memory system scaling challenges is a primary motivation for the exploration of new memory devices with new storage physics and new functionality attributes. The hope is that advances in the ability to manipulate materials at the atomic level can enable the design of a new memory

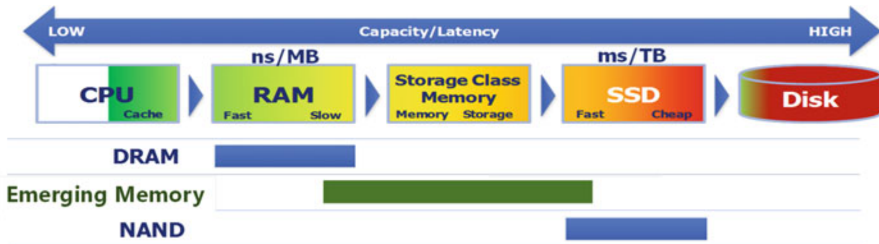


Fig. 11.3 Conceptual memory hierarchy for modern computing system, highlighting the performance and density gap between DRAM and NAND and the possible emerging memory placement [3]

physics device that can mitigate the scaling issues through supplementation of the existing NAND and DRAM and possibly their eventual replacement.

Ideally the elusive "universal memory" would be engineered combining the best of DRAM with the best of NAND. Unfortunately, no such memory physics has been discovered and may be precluded from existing based on the competing requirements of nonvolatility and power or performance and cost. Nonvolatility requires an energy barrier to prevent the loss of data; this energy barrier requires energy to overcome for the write operation. Performance requires short transient time for signals to move across the memory chip often resulting in more circuitry and smaller blocks of memory resulting in higher cost.

An overview of the memory hierarchy today (Fig. 11.3) illustrates several orders of magnitude gap in both performance and density between DRAM and NAND, and filling this gap has been a focus of the new memory technology, often referred to as emerging memory. The memory hierarchy starts with the very fast but low-density caches on the CPU and moves outward to fast and higher-density DRAM potentially in the same package as the CPU, further outward to slow and much higher-density DRAM separate from the CPU. All of these memories are directly addressable by the CPU and all are volatile. On the other end of the memory hierarchy are storage devices where the memory locations are not directly addressable by the CPU. The access latency of these storage devices is large, but their density is also large. NAND has emerged as the primary solid-state storage memory. A gap of around three orders of magnitude for both density and latency exists between DRAM and NAND. This large gap is the target of the emerging memories with potential features as shown in Fig. 11.4.

Consider a representative system of "today" using DRAM as the main working memory and NAND as the storage memory, perhaps a cellular phone, a schematic representation of which is shown in Fig. 11.5a. When the system is started from an off state, it must move all of the code and data required for it to run from the NAND (nonvolatile storage) to the DRAM (volatile main memory). The DRAM and NAND memory will impact the system performance and power.

1. Because the NAND is slow and has a limitation on the number of times it can be used, it is not possible for the system to execute its programs out of the NAND.

	DRAM	Emerging	MLC NAND
Aerial Density – GB/mm^2	1x	10x	30x
Alterability Granularity	2kB Page	16-256B	1MB Block
Read Latency / Bandwidth	20ns / 3,200MT/s	100ns / 2,400MT/s	50us / 667MT/s
Write Completion Time / Bandwidth	20ns / 3,200MT/s	1us / 200MT/s	1ms / 667MT/s
Endurance	"unlimited"	1.E+07	1.E+04
Retention	64ms	1yr	1yr
Management Required	Very Low	Medium	High

Fig. 11.4 Potential emerging memory features to be placed between DRAM and NAND on the memory hierarchy

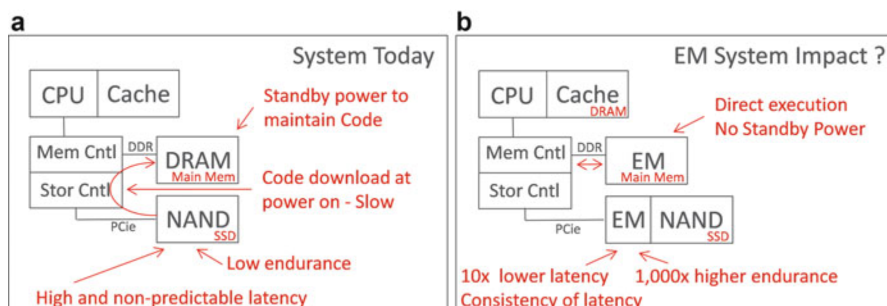


Fig. 11.5 (**a**, **b**) Potential system impact of insertion of emerging memory into the memory hierarchy

The required code and data must be moved to a fast memory, the DRAM. This power-up operation can take considerable time during which the system is not available for use and is nonresponsive to the user.

2. Because the DRAM is volatile, it is necessary to refresh the data it contains at a sufficiently fast rate to avoid data loss. This refresh operation consumes power and must be done regardless of if the DRAM is being used or not. It is not possible to turn the DRAM off to save power even when the system is not in use and preserve a fast wake-up time, often referred to as instant on.
3. Because the DRAM is volatile, the system must insure that the data it contains is not lost if the power is interrupted. To accomplish this, the data is frequently written out to the nonvolatile NAND memory so that it is possible to recover from a power loss event. The mechanisms to do this are commonly called journaling and checkpointing, and they consume power and performance.
4. The NAND memory is slow, and to maintain adequate reliability and maximize the number of times it can be used (called endurance), it requires a large degree of system management. This management results in a further reduction in the performance and makes the accesses to the NAND nondeterministic. Accesses are generally slow but are sometimes very slow due to management. This

nondeterminism can have an impact on systems, since it can stall the process that is running, while it waits on the memory.

5. The NAND memory can wear out with usage. It is only possible to use the NAND memory bits for a fixed number of times; they have a limited endurance. Once the endurance is exceeded, the reliability of the NAND will degrade to the point that it is not useable in the system. A NAND solid-state drive is typically specified by the number of total writes it can achieve.

Now consider how an emerging memory might help to address some of these issues, the potential EM system impact as shown in Fig. 11.5b. The EM does not replace the DRAM or the NAND but instead tries to supplement these memories and reduce their undesirable system traits.

1. The EM is nonvolatile and fast enough to have the potential for the system to execute its programs out of the EM. DRAM is still in the system, but it is much less DRAM and used more as part of the cache hierarchy of the CPU. Since little volatile memory is in the system, it is now possible for the system to power up very quickly improving the responsiveness to the user.
2. Since the EM is nonvolatile, it does not need to be refreshed and can be powered off when it is not being used. This saves power for the system and provides for an instant on feature for the system coming out of a deep standby state.
3. Since the EM is nonvolatile, new procedures for journaling and checkpointing can be realized since data is preserved on power loss. These procedures may still be required to address other system needs, but the implementation can be different to save performance and power.
4. The EM can be used as a cache in front of the NAND SSD in order to improve the responsiveness of the SSD and reduce the variability of the latency.
5. The high endurance capability of the EM can buffer the NAND from frequent writes and reads, improving the responsiveness and reliability of the SSD.

In addition to the above, the higher aerial density of the EM than DRAM can allow for many more bits to be close to the CPU which can result in enabling the CPU to work quickly on larger data sets. This example is relevant for devices ranging in size from servers down to phones. For very small devices, such as a watch or an IOT device, it may be the case that the new memory could be the only memory in the system, reducing the system complexity and form factor.

For additional information related to memory hierarchy, refer to references [5–9].

11.4 The Case for PCM as the First True Emerging Memory

The concerns of technology and system scaling complexity were the motivation for the development of PCM beginning in the late 1990s. Work on the underlying physics of the chalcogenide materials can be traced back far earlier than this

timeframe and was the basis for the determination of chalcogenide as a viable memory material candidate.

Initial product development was focused on the replacement of NOR flash memories with products demonstrated by both Micron and Samsung at near state-of-the-art lithography (~45 nm) and densities (~1Gb) at the time of introduction (~2009). The feature set and main market focus for these devices were targeted at replacing NOR in cellular phones. The product specifications were superior to NOR and were specifically targeted for integration with an LPDDR DRAM. JEDEC standards were developed for the PCM technology reliability, and a common product specification was developed by Samsung and Intel for the LPDDR product. Tens of millions of these devices were used in phones, often stacked in the same package with a LPDRAM. These phones exhibited the expected features enabled by PCM. It is believed that this represents the largest volume of bits ever shipped on a non-charge-based solid-state memory physics. The Micron 1Gb product is shown in Fig. 11.6.

The choice of targeting the NOR cellular market may not have been the best choice. With the advent of "smartphones," the cellular phone architecture started to rapidly move from a NOR-based "execute in place" architecture to a NAND-based "store and download" architecture. This transition was driven by the rapid increase in storage required by the smartphone which was enabled by the rapid increase in density supplied by NAND. The NOR flash market, and thus the PCM opportunity, in cellular phones has decreased dramatically over the past 15 years.

Solid-state drives (SSDs) based on PCM were also developed and prototyped in systems. This used model is similar to the "EM + NAND" storage shown in Fig. 11.7. Significant performance improvements of the SSD were demonstrated over NAND only-based drives for read latency and bandwidth. PCM-based SSDs were prototyped and extensive industry testing was done, but products were never introduced. The value of PCM for improved SSD performance was clear, but the density was not high enough to justify production.

The initial PCM products did not deliver the full capability of the PCM technology, in particular, as related to write performance, largely because it was not required by the initial NOR compatible end products. The specifications of the initial products were close to the targets (Fig. 11.4) for read performance, endurance, and nonvolatility but were significantly behind the targets for aerial density and write performance. The second-generation PCM products were targeted to realize the full performance capability of the technology (Fig. 11.8). While these product specs were getting close to the goals for performance and endurance, a significant gap still remained for the density (cost) as compared to DRAM. Ultimately the cost was too high for the feature set delivered, and while the technology was successful in manufacturability, endurance, and performance, the density was not high enough (cost not low enough) to be consistent with the feature set. The technology was not capable of delivering an "overwhelming compelling value" adequate to get the systems to make the changes required for its adoption.

For additional information related to PCM products, refer to references [7, 9–17].

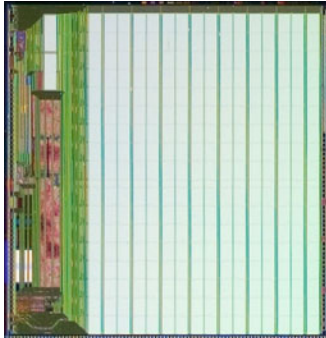


Fig. 11.6 Micron 1Gb phase change memory for NOR flash replacement in cellular phones [10]

		PCM Prototype SSD	NAND NAND SSD
System Interface		PCIe 3.0x8	PCIe 2.0x8
Sector Access Granularity		4096 byte	4096 byte
Random Reads	Bandwidth	5.5 GB/s > 1M IOPS	3.0 GB/s 725,000 IOPS
	HW Latency	5µs	~50µs
	SW Latency	TBD	
Random Writes	Bandwidth	~625 MB/s	400 MB/s
	HW Latency	≤ ~204 µs	300 µs
	SW Latency	TBD	
Density		64 GByte	350GB, 700GB
Total Pbytes (Life)		≥ 120 PB	25 PB (350GB)

Fig. 11.7 PCM SSD prototype

	DRAM	2nd Gen PCM	Emerging	MLC NAND
Aerial Density – GB/mm²	1x	1x	10x	30x
Alterability Granularity	2kB Page	32B	16-256B	1MB Block
Read Latency / Bandwidth	20ns / 3,200MT/s	80ns / 1,066MT/s	100ns / 2,400MT/s	50us / 667MT/s
Write Completion Time / Bandwidth	20ns / 3,200MT/s	1us / 150MT/s	1us / 200MT/s	1ms / 667MT/s
Endurance	"unlimited"	1.E+07	1.E+07	1.E+04
Retention	64ms	1yr	1yr	1yr
Management Required	Very Low	Medium	Medium	High

Fig. 11.8 Anticipated second-generation PCM product specifications

11.5 The Need for "Overwhelming Compelling Value"

Current system hardware and software architectures are heavily entrenched with many years of evolved development built into them in order to exploit the capabilities and mitigate the deficiencies of the existing DRAM and NAND memories. An emerging memory, such as PCM, must provide an "overwhelming compelling value" to justify the system hardware and software changes required to use the memory. The larger the

required system changes, the larger must be the compelling value. This value will be a combination of nonvolatility, performance, and density (cost).

Possible "overwhelming compelling values" for emerging memory:

"End of Scaling" for DRAM/NAND → Dangerous since >$1B/year is spent on scaling research

Development of new usage models → Optimistic and slow since systems must change

Scaling "enabler" for DRAM/NAND → Mitigate DRAM/NAND issues?

One of the primary compelling values for any memory has always been cost and density. This sets a challenge for PCM since, to achieve low cost, it must be at the same litho generation as the incumbents. This is a difficult challenge since the strategy for an emerging memory to be cost-effective for technology development is typically to draft off of the incumbent technologies putting it, by definition, behind in time. The first-generation PCM products demonstrated the value of nonvolatility and much higher than NAND endurance. The second-generation products, which were not introduced to the market, showed the promise of much improved performance. Density became the largest PCM challenge, a challenge magnified by the transition of NAND to 3D structures. PCM needs a path to lower cost, most likely through 3D structures.

An alternate, or supplemental, path to 3D structures for lower cost is storing multiple bits in a single memory cell as is done today on flash memory. PCM has demonstrated this ability as reported by several companies; one such effort is shown in Fig. 11.9. Storing multiple bits in a single cell is more complex than storing single bits and usually requires a mature, stable technology. The algorithms used to write and read multiple bits in a single cell result in slower write and read performance which reduces the performance value that PCM provides. Multi-bit per cell storage may provide a path to lower cost for certain applications that can tolerate the performance impact.

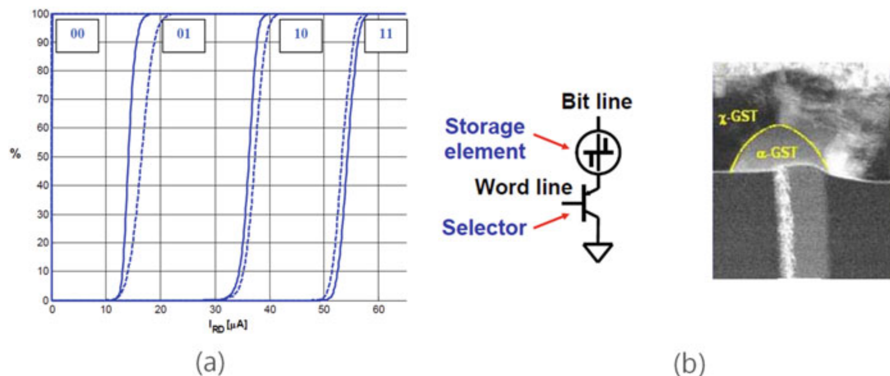


Fig. 11.9 Two bit per cell storage in a 256 Mb PCM device. Figure (**a**) shows array distributions of the four distinct states that make up 2 bits. Figure (**b**) shows the memory cell structure [18]

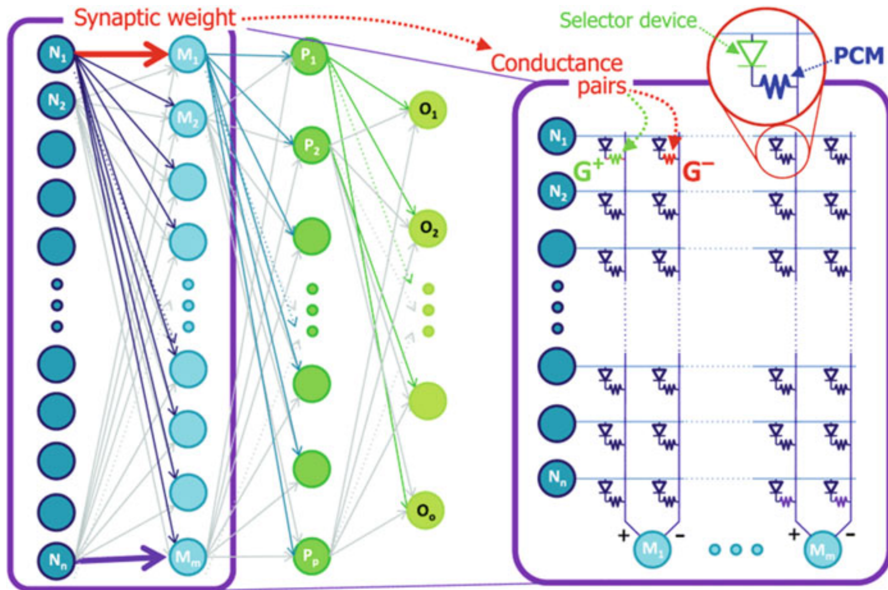


Fig. 11.10 Neuro-inspired computing, in which neurons activate each other through dense networks of programmable synaptic weights, can be implemented using dense crossbar arrays of nonvolatile memory (NVM) and selector device pairs [19]

The unique “summing” and “thresholding” properties of the phase change materials make them a good candidate for neuromorphic or “brain-inspired” computing systems. A silicon-based system with PCM used as a synapse in a three-layer perceptron network (Fig. 11.10) of 164,885 synapses was trained on a database of 5000 handwritten digits with accuracy of 82% [19].

If classification accuracies can be improved, these neural network systems have the potential for 10× faster execution and 100× lower power compared to conventional Von Neumann computation.

For additional information related to compelling value, refer to references [18–20].

11.6 The Challenge for PCM in a “3D” Memory World

The same thing that drove the investigation of PCM, the scaling issues of DRAM and NAND, is also becoming the challenge for PCM as the incumbent technologies move to 3D integration to address scaling. To effectively compete with the incumbents for cost, PCM will also need to move to 3D. Several demonstrations of 3D PCM have been published in the technical literature (Fig. 11.11). These technologies use a thin film selector device to replace the silicon substrate-based selector used in the PCM technologies. This allows for 3D integration, placing multiple

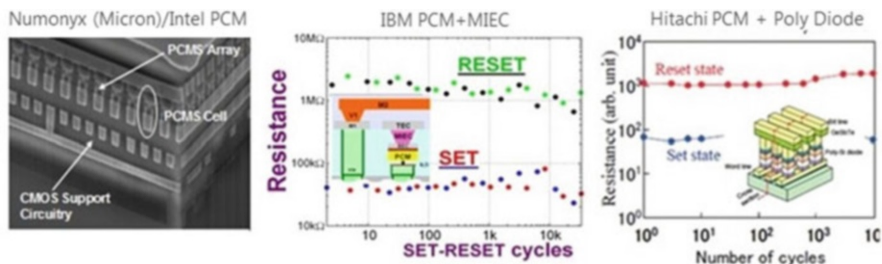


Fig. 11.11 3D PCM demonstrations [21–23]

layers of cells over circuitry in the silicon substrate and significantly increasing the storage density per square cm of silicon area. With NAND moving to 3D, it now appears that PCM (or any other emerging memory) will need to enter the market as a 3D technology in order to provide the required density and cost.

For additional information related to 3D PCM, refer to references [21–25].

11.7 Summary

Memory is becoming central to performance and power in modern compute systems, often defining the capability of the system as much as the CPU. Computer systems have evolved to adapt to the idiosyncrasies of the dominant memory technologies of DRAM and NAND. Emerging memories have the promise to supplement DRAM and NAND in systems providing some level of mitigation for their weakness, those being DRAM volatility and NAND reliability. Possibly more importantly, the emerging memories provide an opportunity to bridge the large performance and density gap between DRAM and NAND, inserting a new level of hierarchy and moving more "fast and nonvolatile" bits temporally close to the CPU.

Phase change memory emerged as the first of the new memory types with the promise to combine features of DRAM and NAND. PCM was initially targeted at replacing NOR flash in cellular phones, a goal which was achieved in production to a limited degree. The early PCM product demonstrators showed the value of the emerging memory features for both memory-mapped (phone) and storage-mapped (SSD) usage models either through products or product-like prototypes. The PCM SSD prototypes delivered the world's fastest SSD in 2014, and tens of millions of phones with PCM were shipped. Research labs have demonstrated the ability to store multiple bits in a single PCM cell showing a potential path to lower cost as well as the use of the PCM cell in neural networks showing a path to potential new usage models.

Ultimately the PCM problem became one of cost. The cost structure was not consistent with the enhanced features. This points to the need for PCM to focus on lower-cost architectures, such as 3D architectures. Several 3D architectures have been demonstrated in conference papers, pointing to the future for PCM evolution.

References

1. K. Prall, Benchmarking and metrics for emerging memory, IMW, 2017
2. G. Atwood, Current and Emerging Memory Technology Landscape, Flash Memory Summit, 2011
3. S. DeBoer et al., A semiconductor memory development and manufacturing perspective, ESSDERC, 2014
4. S. Lee, Technology scaling challenges and opportunities of memory devices, IEDM, 2016
5. S. Eilert et al., Phase change memory: a new memory enables new memory usage models, IMW, 2009
6. Rich. Freitas, Storage class memory: technology, system, and applications, in *IEEE Hot Chips 22 Symposium*, 2010
7. A. Akel, Onyx: a prototype phase change memory storage array, in *HotStorage'11 Proceedings of the 3rd USENIX conference on hot topics in storage and file systems*, 2011
8. Next Generation Memory. IEEE Comput. Soc. 46(8), 21–59 (2013)
9. H. Kim, Evaluating phase change memory for enterprise storage systems: a study of caching and tiering approaches, in *FAST'14 Proceedings of the 12th USENIX conference on File and Storage Technologies*, 2014
10. G. Servalli, A 45nm generation phase change memory technology, IEDM, 2009
11. F. Pellizzer et al., Novel uTrench PCM cell for embedded and stand-alone NVM applications, in *Symposium on VLSI Technology*, 2004
12. F. Bedeschi et al., An 8Mb demonstrator for high-density 1.8V phase-change memories, in *Symposium on VLSI Circuits*, 2004
13. F. Pellizzer et al., A 90nm phase-change memory technology for stand-alone non-volatile memory applications, in *Symposium on VLSI Technology*, 2006
14. H.G. Lee, High-performance NAND and PRAM hybrid storage design for consumer electronics. IEEE Trans. Consum. Electron. **56**(1), 112–118 (2010)
15. H. Chung et al., A 58nm 1.8v 1Gb PRAM with 6.4MB/s program BW, ISSCC, 2011
16. Y. Choi et al., A 20nm 1.8v 8Gb PRAM with 40MB/s program bandwidth, ISSCC, 2012
17. HGST, HGST Research Demonstrates Breakthrough Persistent Memory Fabric, Flash Memory Summit, 2015
18. F. Bedeschi et al., A multi-level-cell bipolar-selected phase-change memory, ISSCC, 2008
19. G. Burr et al., Recent progress in phase-change memory technology. IEEE J. Emerging Sel. Top. Circuits Syst. **6**, 146–162 (2016)
20. M. Stanisavijevic et al., Demonstration of reliable triple-level-cell (TLC) phase change memory, IMW, 2016
21. D. Kau et al., A stackable cross point phase change memory, IEDM, 2009
22. Y. Sasago, Cross-point phase change memory with 4F2 cell size driven by low-contact-resistivity poly-Si diode, in *Symposium on VLSI Technology*, 2009
23. K. Gopalakrishnan, Highly-scalable novel access device based on mixed ionic electronic conduction (MIEC) materials for high density phase change memory (PCM) arrays, in *Symposium on VLSI Technology*, 2010
24. K. Lee et al., A 90nm 1.8v 512Mb diode-switch PRAM with 266MB/s read throughput, ISSCC, 2007
25. D. Kang et al., Considerations on highly scalable and easily stackable phase change memory cell array for low-cost and high-performance applications, NVMTS, 2014

Index

A
Ab initio molecular dynamics (AIMD), 136
Ag-/In-doped Sb_2Te alloys (AIST alloys), 126, 153
Aggressor, 114
Amorphization current densities, 231
Amorphous as-deposited state (A-AD), 199
Array biasing
- deselect BL voltage, 295, 296
- dummy line, 297
- implementation, 295
- line coupling, 298
- power/energy considerations, 299, 300
- pull-down, 296
- pull-up, 295
- simplest row decoding approach, 297
- single cell, 295
- standby, 299
- tiles, 295–298
- voltage compensation, 298
- voltage *vs*. current, 299

Arrhenius model, 93
Atomic force microscope (AFM), 250

B
BER tolerance *vs*. time, 122
Bipolar junction transistors (BJTs), 3, 17
Bit error rate (BER), 120, 121
Bond angle distribution, 145
Bottom electrical contact (BEC), 55, 244
Brain-inspired computing systems, 322
Built-in heater device architectures
- adhesion layer, 274
- cross-spacer structure, 279, 280
- dash-confined structure, 278, 279
- interfacial layer, 272, 273
- matchstick structure, 270–273
- materials, 270
- phase transition, 269
- pore structure, 270–272
- ring structure, 274, 275
- μTrench structure, 275–277
- wall structure, 276, 277

C
Carbon, 194–200
Carbon nanotube (CNT), 250, 281, 282
Cell temperature (T_2) *vs*. cell dimensional scaling (F), 116
Chalcogenide alloys
- *Angell* plot, 52
- *DZ** values, 61
- fragility behavior, 51–54
- Kissinger plot, 52, 53
- number of migrating atoms, 57
- polarity pulse, 59
- segregation behavior, 54–62
- SEM images, 60
- velocity/viscosity decoupling exponent and fragility behavior, 54

Chalcogenide composition *vs*. distance, 108
Chalcogenide materials
- amorphous and polycrystalline phase, 126, 127
- automotive applications, 131
- compositions, 130
- computer simulations, 172
- crystallization mechanism, 129, 172

© Springer International Publishing AG 2018
A. Redaelli (ed.), *Phase Change Memory*,
https://doi.org/10.1007/978-3-319-69053-7

Chalcogenide materials (*cont.*)
  device performances and optimization, phase change material, 131
  electrical field, 126
  electropositive elements, 129
  integration process flow, 131
  memory applications, 127
  mushroom-type cell, 128, 130
  optimizing memory cell design, 132
  prototypical alloys, 132
  resistance drift, 132
  resistivity and reflectivity measurement, 127
  resonant bonding, 128
  scaling effects, 132
  ternary phase diagram, 128
Chemical mechanical polishing (CMP), 270, 276
Chemical vapor deposition (CVD), 278
Classical nucleation theory (CNT), 187–189, 251
Coordination numbers (CN), 144
Crystalline GST (xGST), 4
Crystallization temperature T_c, 191
Current-voltage (I-V) curves, 18, 19, 21
Cycle lifetime, 50
Cycling failure distribution, 107

D
Data retention
  automotive applications, 96
  bake retention testing, 95
  bake-to-bake consistency, 94
  Boltzmann's constant, 93
  cell resistance distributions, 92, 93
  cell-to-cell structure, 93
  chalcogenide material, 95
  crystallization mechanism, 94, 95
  dopants, 98, 99
  GeSbTe system, 92
  lognormal distribution, 93
  MLC, 100
  nucleation and growth times, 95
  PCM resistance *vs*. time curve, 92
  programming voltages, 96
  RESET resistance, 92
  resistance drift, 92
  stoichiometry/doping, 97
  thermal activation energy, 100
Density functional (DF), 136
Density of states (DoS), 165
Deposition method, 169
Differential scanning calorimetry (DSC), 52, 133
Diffractive coherent domain (DCD), 160
Diodes, 3
3D Memory World, 322, 323
Drift phenomenon, 21

E
Electrical resistivity, 183
Electromigration, 57
Electrostatic storage physics, 314
Electrothermal model, 41–43
Elemental analysis, 55
Elemental segregation, 104
Emerging memory (EM), 316–319
E_g^{04} method, 149
Energy dispersive X-ray (EDX), 56
Energy landscape, 22, 23
Equilibrium Fermi energy, 13
Equilibrium phase diagram, 203
Error-correcting code (ECC), 121
Extended X-ray absorption fine structure (EXAFS), 135, 165

F
Field-effect transistor (FET), 17, 18
Finite element software (COMSOL™), 240
Fourier equation, 42
Fourier's law. *See* Thermal resistance
Fourier-transform infrared (FTIR) spectroscopy, 138
FTIR absorbance spectra, 157–159

G
$Ge_2Sb_2Te_5$ and GeTe alloys
  amorphous phases, 142–146
  crystalline phases, 139–142
  methods, 135–138
  transport and optical properties, 148–151
  volume change upon crystallization, 146
Germanium nitride (GeN_x), 159
GeSbTe (GST), 186–213
  alloys, 132–134
  automotive applications, 182
  crystalline phase transformation, 181
  device performance and reliability, 182–186, 213
  dopant element, 214
  high-capacity, 181

nonstoichiometric (*see* Nonstoichiometric alloys)
nonvolatile memories, 181
PCMs, 181
phase diagram, 126
resistivity, 217
SET and RESET resistance distributions, 216, 218
SET programming pulses, 216
stoichiometric (*see* Stoichiometric compounds)
thermal stability, 182
TiN ring bottom electrodes, 216
GeTe rhombohedral unit cell, 140
Gibbs free energy, 241
Gillespie algorithm, 241
Gillespie Cellular Automata (GCA), 240
Growth-dominated materials, 153

H
Hard x-ray photoemission spectroscopy (HXPES), 13
Heat confinement, 47
Heater, 17
Heterogeneous nucleation, 171
Historical development, PCM
amorphous chalcogenide alloys, 2
commercial volume production, 3
dry photocopying technique, 2
GST, 3
photoconductive properties, 2
solid-state memory devices, 3
solid-state NVM device, 2
stages, 4
technology viability, 3

I
In situ self-anneal (ISSA), 109
Interface engineering, 172
Ion migration model, 61
Isotropic scaling, 115, 255

J
Joule effect, 41
Joule heating, 17, 42

K
Kissinger analysis, 228

L
Laser-induced reversible phase transition, 3
Localized states, 11, 12, 17, 24
Low-power double-data-rate nonvolatile memory (LPDDR2-N), 4

M
Maxwell constraint counting method, 157
Memory
applications, 321
DRAM and NAND, 316–318
EM system, 317, 318
endurance, 317
functionality attributes, 315
hierarchy, 316
high-end systems, 313, 314
modern computing systems, 313, 314
NAND and DRAM, 315
nonvolatility, 316
scaling and system challenges, 314, 315
storage devices, 316
switching, 6
Metal–oxide–semiconductor field-effect transistors (MOSFETs), 3
Micron technology, 119
Mixed ionic-electronic conduction (MIEC), 17
Molecular beam epitaxy (MBE), 142
Molecular dynamics (MD), 136
Moore's Law, 223
Multilevel cell (MLC), 100, 161, 269
Mushroom-type PCM cells
device operating parameters, 236
electrical, thermal and phase transformation, 240
events, 241
experimental scaling, 242, 243
GCA, 241
geometrical and physical parameters, 236
heater contact diameters, 241
isotropic scaling, 236, 239
melting voltage, 238
ohmic representation, 236
optimum cell, 240
parasitic heating, 240
phase-change (GST) layer, 241
scaling behaviours, 236
technology node, 238, 239
thermal conductivity, 237
thermal crosstalk, 240, 242
thermal resistance, 237

N
Negative differential resistance (NDR), 18
Neuro-inspired computing, 322
Nitrogen, 189–193
Non-isotropic scaling, 115, 255
Nonstoichiometric alloys
  conductive path, 209
  drift coefficient, 205
  Ge-rich GST, 202–212
  GeSbTe material engineering, 202
  low-resistance states, 206
  programming cartographies, 210, 211
  Sb-rich GST, 212–214
  SET programming, 208
  STEM/EELS images, 210
  Te-poor GST, 213, 215
  transmission electron microscopy, 207
Nonvolatile memory (NVM), 129, 322

O
Octahedral environment, 144
Overwhelming compelling value, 320–322
Ovonic threshold switch (OTS), 17, 247
Oxygen, 200

P
Pair distribution function (PDF), 135
Parasitic capacitance, 30
Partial pair distribution functions (PDF), 137
PCIe-attached storage arrays, 7
PCM array, 303–307
  biasing voltages, 289
  chip yield, 292, 294
  column-row short, 293
  diode selectors, 286
  ECC, 308, 309
  embedded firmware, 308
  fuses, 309, 310
  global word line decoder, 289, 290
  implementation, 285, 286
  microcontroller, 308
  multilevel cells, 307
  parameters, 285
  partitions, 288, 289
  program modes, 289
  program operation
    blind program, 305
    buffer load, 303
    erase operation, 306
    flowchart, 303, 304
    loops, 303
    methodologies, 303
    microcontroller, 303
    program buffer, 306
    program calculator tool, 304
    program disturbs, 305
    pulse duration, 304
    suspend commands, 306, 307
  read while write/banking, 290, 291
  redundancy, 291–293
  selected and deselected cells, 285, 286
  sense amp hierarchy, 290
  sense amplifiers and program loads, 285
  spare tiles, 294
  tile sizing, 287
  word and bit lines, 285
PCM scaling
  AIST, 226
  amorphization process, 224
  binary, ternary/quaternary chalcogenide alloys, 225
  compositions, 226
  Ge-Te-Sb ternary phase diagram, 225
  one dimension, BNF–229
  parameters, 224
  performance metrics, 226
  phase-change superlattice materials, 226
  silicon-based Flash and DRAM, 223
  technology node, 223, 224
  three dimension, 232–234
  two dimension, 229–231
PCM SSD prototype, 320, 323
Peierls (symmetry breaking) distortion, 139, 150
Phase change materials
  ab initio simulations, 156
  amorphous phase, 155
  boson peak, 157
  carbon environment, 156
  carbon segregation, 160
  C-doped and N-doped alloys, 155, 159
  crystallite growth process, 152
  crystallization temperature and mechanism, 168–172
  dopant elements, 155
  engineering, 252–255
  film deposition method, 156
  high-resolution transmission electron microscopy, 160
  issues, 166–168
  nanosecond timescale and nanometer length scale, 153
  nucleation-dominated/growth-dominated, 151, 153

retention time, 155
tetrahedral and defective octahedral environments, 151
thermal conductivity, 161
Phase-change memories (PCMs). *See* Chalcogenide materials
attributes, 6
device performance and applications, 6, 7
prototype, 5
working principle and device physics, 4–6
Phase-dependent band structure
electrical characterization, 14–17
electrical resistivity, 11
GST, 11, 12
optical characterization, 12–14
photon absorption/reflection, 11
thermodynamic stability, 11
Photothermal radiometry, 195
Poole-Frenkel (PF) conduction, 19–23
Power consumption, 7, 41, 42, 62
Power spectral density (PSD), 31
Probabilistic cellular automata (PCA), 241
Product-oriented PCM cell designs
crossbar cells, 246, 247
pore-type cell, 245, 246
μTrench and dash-type cell, 244, 245
Product-oriented *vs*. research-oriented PCM cells, 257
Programming currents, 250, 252, 255
Programming current *vs*. resistance curve, 243

R
Radiation sensitivity/soft errors, 119, 120
Rapid nucleation materials, 187
Raw bit error rate (RBER), 122
Read disturb, BNF–113
Read latency, 41
Reflow soldering temperature profile (RSTP), 206, 207
Reliability, 92–100
cycling endurance, BNF–109
data retention (*see* Data retention)
DRAM and NAND, 89
error management, 121–123
failure modes and mechanisms, 90, 91
high-radiation environments, 90
memory cells, 89
nonvolatile memory applications, 89
radiation sensitivity/soft errors, 119, 120
read and write latency, 89, 90
read disturb, 110–114
thermal disturb, 114–119
Remote heater device architectures
aggressive memory cell footprint, 280
dielectric layer, 280
micro-heater structure, 280, 281
Research-oriented PCM cell designs
CNT-based cells, 249–251
lateral, line and bridge cells, 248, 249
physical and electrical behaviours, 248
Resistance drift phenomenon, 32, 33, 92, 132, 161, 173, 184
Resistance–power ($R-P$) curves, 43, 44
Resonance bonding, 128, 151, 183
Retention time, 155
Rocksalt cubic unit cell, 134

S
Scanning transmission electron microscopy (STEM), 167, 169, 170
Scientific community, 8
Self-consistent environment, 8
Self-heating device architectures
bridge structure, 267, 268
chalcogenide materials, 263
Joule heating, 263
line structure, 266, 267
metallic liner structure, 269
pillar structure, 263–265
R-I curves, 264
SHW, 265, 266
via structure, 268, 269
Self-heating wall (SHW), 265, 266
Semiconductor devices, 8
Sense amplifier, 300–303
SET-MIN state, 199
SET resistance, 42
Short-range order (SRO), 184
Single event upset (SEU), 119
Single-nanometre technology nodes, 256
Slow nucleating materials, 187
Solid-state drives (SSDs), 9, 319
Solid-state storage subsystem, 7
S-shaped negative differential conductance (SNDC), 5
Stability. *See* Phase change materials
Stoichiometric compounds
carbon, 194–200
composites materials, 200, 201
crystallization process, 189
data retention, 193
electrical and thermal resistivity, 189
$Ge_2Sb_2Te_5$ and GeTe, 186–188, 190
GeTeN alloys, 192

Stoichiometric compounds (*cont.*)
  Kissinger representation, 189
  nitrogen, 189–193
  nucleation and growth rates, 189
  oxygen, 200
Storage-class memory (SCM), 4, 6, 8
Structural relaxation and resistance drift, 32, 33
  aging, 162, 163
  coordination shells, 165
  drift exponents, 163
  EXAFS, 165
  Ge-Ge bonds, 164
  incubation time, 162
  material conductivity, 164
  MLC, 161
  multilevel storage devices, 161
  physical mechanisms, 163
  power law, 162
  resistivity (logarithmic scale), 162
  structural characterizations, 164
  tetrahedral environments, 165
  thermal annealing, 164
Subthreshold slope (STS), 21
Superlattice-like (SLL), 253
Switching dynamics, 28–31
Switching energy, 255
Switching power, 255
Switching speed, 256
Synchrotron radiation, 135
System-on-chip/embedded application, 3

T
Tauc-Lorentz law, 150
Temperature–voltage ($T-V$) diagram, 46, 47
Tetrahedral environment, 144
Thermal boundary resistance (TBR), 117, 118, 229
Thermal disturb, BNF–119
Thermal engineering, 251, 252
Thermal resistance
  current/power variation, 49
  GeSbTe alloy, 49, 50
  PCM cell geometry, 47, 48
  room temperature–power plot, 50
  and temperature estimation, 43–46
Threshold switching, 5, 23–27, 35
Threshold voltage V_T, 18
Time-dependent transport, 32–34
Top electrode contact (TEC), 48
Transmission electron microscopy (TEM), 17, 22, 58, 130

U
Ultimate scaling, 249–251
Umbrella flipping model, 151
Universal memory, 316

V
Vapour-liquid-solid (VLS), 229
Vibrational density of states (VDoS), 136, 138, 157

W
Wireless systems, 6
Word-line plug (WL plug), 264

X
Xerography technology, 2
X-ray absorption near-edge structure (XANES), 136
X-ray absorption spectroscopy (XAS), 135
X-ray diffraction (XRD), 139, 160, 226, 232