Università degli Studi
di Modena e Reggio Emilia

DOTTORATO DI RICERCA IN FISICA E NANOSCIENZE

XXVI CICLO

ELECTRICAL TRANSPORT IN CHALCOGENIDES AS MATERIALS FOR PHASE-CHANGE MEMORIES

Candidato: Andrea Cappelli

Relatore: Prof. Carlo Jacoboni

Tesi per il Conseguimento del Titolo di
Dottore di Ricerca

Direttore della Scuola di Dottorato: Prof. Franca Manghi

Febbraio 2014

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

One of the main challenges that the computing industry is facing nowadays revolves around memory issues. The ability to store, modify and retrieve data is becoming at least as critical as the power to process them, particularly as processing units become increasingly fast, making the memory elements the real bottleneck for computation.
Present technology relies mainly on Dynamic Random Access Memories (DRAM). DRAMs provide great speed but relatively poor storing capacity, besides being volatile devices (i.e. they lose the stored information after turning off the whole computing system).
A proper candidate to the role of new-generation memory must prove competitive with respect to the present technology according to several criteria: the new device must be non-volatile, have good speed in both writing and reading operations, exhibit low power consumption and show promising scalability outlooks in terms of memory capacity per unit space, all of these properties ideally being complemented by low production costs.
Two solid-state technologies that match the requirements described above are Flash and Phase-Change Material (PCM) memories. Flash memories exploit the charge trapped in the floating gate of a transistor to encode the bit of information to be stored.
PCM technology makes use of the controllable and reversible transition between amorphous and crystalline structures that can be induced in a nanometric sample of chalcogenide material, typically a GeTe–Sb_2Te_3 alloy. One of the most promising is the $Ge_2Sb_2Te_5$ compound (GST for brevity). Two architectures are being studied, one of them employing PCM devices as a replacement for DRAM themselves, the other using PCM devices as an intermediate layer between mass storage units and DRAM [117].

Chalcogenide materials have already been used to fabricate optical storage devices,

such as rewritable DVDs: in that case, the markedly different reflectivities of the two phases of the compound have been exploited to distinguish the amorphous or crystalline state of the bit. During the writing operation, a highly focused laser beam was employed to heat the bit region, followed by a quick or gradual cooling step that left the bit in the amorphous or crystalline state, respectively. The reading capability is then achieved through a low-power laser beam, unable to induce any change in the chalcogenide material.

The use of PCM as electrical non-volatile random access memories (PCRAM) follows an analogous approach, based on the very different electrical properties of the two structural phases, and some specific features of their current-voltage (IV) characteristic curves, as depicted in Figure 1.1. The crystalline phase shows a mostly Ohmic behavior,

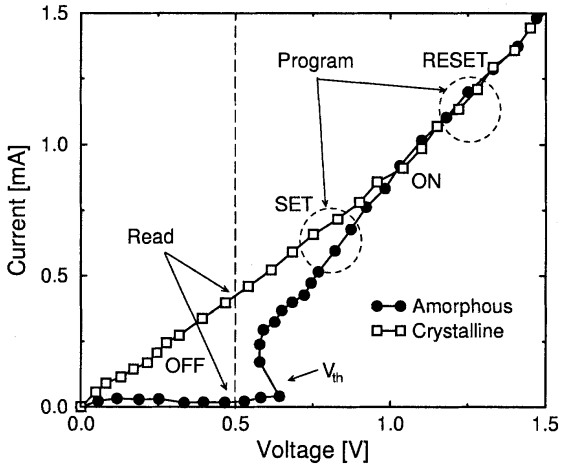

Figure 1.1: Current-voltage characteristic curve for the crystalline (white squares) and amorphous (black circles) phases of GST. The Read operation voltage is shown in the sub-threshold region, together with the Programming point for SET (crystallization) and RESET (amorphization) operations above the threshold voltage V_{th}. Figure taken from [93].

with a quite high conductivity.

The amorphous phase IV, on the contrary, exhibits a very low-conductivity Ohmic behavior at low applied voltages; for high-voltages, the curve is instead characterized by an S-shaped Negative Differential Conductivity (S-NDC) region, which then leads to a high-conductivity branch for even higher voltages. This implies that when a bias greater than a critical threshold voltage V_{th} is applied, the resulting current I in the device belongs to the upper high-conductivity branch of the IV curve. Thus, for high enough voltages, a high current can be obtained although the device is still amorphous. This phenomenon is known as Ovonic Threshold Switching (OTS), and it is different from the

phase-change switching between the two structural states.
These electrical properties can be employed for memory device operations. Instead of a laser, electrical pulses are used to operate the device. A short low-voltage pulse allows to distinguish the resistance of the bit (marked by READ in Figure 1.1), and thus its structural state and logical value. Instead, a higher-voltage signal is used to alter the state of the memory cell via Joule heating: amorphization is obtained from the crystalline state by melting the memory cell with a high-current pulse and subsequently quenching it into the disordered structure with a very fast cooling. Crystallization starting from the amorphous phase relies on Joule heating as well, and is strongly helped by the OTS phenomenon: a bias slightly beyond V_{th} allows to reach currents which are high enough to start a structural rearrangement, then a slow cooling step lets the structure settle in an ordered lattice.

Besides memory applications, these materials are also proposed to implement beyond-Von-Neumann computing units and neuromorphic or cognitive applications [127, 62].
It is clear that understanding the electrical properties of phase-change materials, and particularly the nature of the S-NDC, is of great importance for a fundamental understanding of the physics of such materials as well as for a better modeling and design of chalcogenide-based devices.

1.2 Aims of the project

The scientific aim of this thesis is the theoretical and experimental investigation of the electrical transport properties of chalcogenide PCM, with emphasis on their switching behavior.

From the theoretical side, the origin of the S-shaped IV curve of amorphous cells was studied employing several models formulated in our group, each of them examining the phenomenon on a different level.
Experimental evidences suggest that localized states in the band gap of the amorphous material play a fundamental role in determining its characteristic transport properties [48]. Thus, a quasi 1D approach was devised at first, using the concepts of charge and energy transport for a population of carriers occupying localized states in the bandgap of the material. These states were assumed as uniformly distributed in both energy and position. A hydrodynamic-like (HD) approach was employed, having similarities with that described in [45], with the addition of the Poisson equation to guarantee the electrostatic self-consistency of the system. The model was used to study the profiles of electric

potential, carrier population and energy along the sample, in a hot-carrier picture.
The HD-like theory, as well as several other analytical approaches in the literature, neglects for the sake of simplicity the role of the device real geometry in the three dimensions, by introducing the one-dimensional uniformity assumption. Although this does not prevent from capturing the S-NDC behavior, other relevant effects are missed.

Next, we devised a 3D model based on a network of randomly placed sites, representing single localized states or groups of them. In analogy with the hot-carrier concept of the HD-like model, with this theory we investigated steady-state conduction in an arbitrarily-shaped chalcogenide sample by considering rate balance equations for charge and energy transport for carriers hosted by the sites, with rates describing the transitions between them according to an emission-travel-recapture process. Electrostatics is accounted for via the finite-element-method (FEM) solution of the Poisson equation. This approach overcomes the uniform 3D limitation, thus allowing for the analysis of the intrinsic variability of macroscopically equivalent devices, as well as the investigation of the OTS onset on a microscopic basis.
A further level of comprehension of the threshold-related phenomena must include the crystallization process in the transport analysis: current-driven IV measurements show a sudden decrease in the device voltage when the threshold conditions are reached, instead of a complete S-NDC curve, as observed, e.g., in GST ultra-scaled devices with nanotube contacts [129]. We extended the network model in order to account for thermal phenomena and the onset of crystallization, coupling the network model scheme with the FEM solution of the Fourier equation for heat transport in the chalcogenide structure.

On the experimental side, we assessed the performances of novel GST-based devices, employing graphene contacts with different structures. Experimental IV characteristics were obtained for GST cells bridging graphene nano-ribbons (GNR) and GST nano-strips contacted by graphene micro-ribbons (GμR). From these data, key parameters such as threshold voltage, current (I_{th}) and contact resistance were extracted and statistically analyzed to judge the effectiveness of the graphene-based solution.

The candidate performed the activities which are the subject of this thesis within and in collaboration with two research groups.
The theoretical investigation took place mostly at the University of Modena and Reggio Emilia, at the Monte Carlo Laboratory, in Prof. Jacoboni's group. The candidate co-developed the quasi-1D hydrodynamic-like approach and took part in its testing and

calibration. As regards the 3D network model, the candidate devised the theoretical concept behind the model and implemented the transport framework in a Fortran code, integrating it with existing in-house FEM libraries for the solution of the Poisson equation. Finally he co-devised the heat-transport extension of the network model and analyzed simulation results.
The experimental activity was conducted at the University of Illinois at Urbana-Champaign, during a nine-months visit in Prof. Pop's group. The candidate was trained on and performed the electrical characterization of non-volatile memory prototypes fabricated by collaborators at the University of Illinois, and analyzed the characterization results.

1.3 Structure of the document

This thesis is organized as follows. Chapter 2 deals with the physics of phase-change materials, presenting the structural and electronic properties of chalcogenides, GST in particular.
In Chapter 3, we review the main experimental evidences and theoretical models used to study charge transport and switching phenomena in amorphous chalcogenides.
The quasi-1D hydrodynamic-like model is presented in Chapter 4. It is employed to study the transport properties of amorphous-phase devices, also taking into account the effect of different density-of-state hypotheses.
Chapter 5 introduces the 3D network model. Its results are presented, with respect to conduction properties and variability analysis. The Ovonic switching mechanism is microscopically studied. The extension of the model to include thermal phenomena and crystallization is eventually described.
Chapter 6 presents a novel kind of graphene-contacted GST memory devices. The fabrication process is explained, and device performances are assessed.
In Chapter 7 some concluding remarks are made, together with a summary of the results achieved in this thesis work and possible future developments or required further research.

Part I

Phase-change materials for memory devices

Chapter 2

Physics of phase-change memories

The ground-breaking work of Stanford R. Ovshinsky [85] is universally regarded as the seminal one for the birth of the whole field of chalcogenides for data storage applications. In his 1968 paper, Ovshinsky described successful memory operations for a Te-As-Si alloy device for several cycles and over a period of several months. This ignited the interest of both academic and industrial communities, partly because of the rich physics in these new materials, but partly (and very importantly) also due to the strong drive of possible applications in the exploding electronics and computer industry of that age.

The importance of this last aspect is even clearer looking back at the 1970's. In this period, the first prototype batches of PCM devices were produced, the first of them being described in an already evolved form in the paper from Neale *et al.* [81]. An array of 256-bit was demonstrated, each of them composed of an amorphous element in series with a $p-n$ junction diode. Another batch was presented by Shanks and Davis in 1978 [110], with a size of 1024 bits and $p-n$ junction diodes replaced by Schottky ones. Unfortunately, both these pioneering efforts showed too high power consumption levels, with μA to mA current needs and $10-20$ V requirements. Thus, they were unable to compete with other contemporary technologies, and ultimately to make it to the market. The interest in these materials as solid state memories cooled down for a couple of decades.

At the end of the 90's, a renewed interest in the subject came about due to big improvements in controlling one fundamental property of all electronic devices: size. Energy consumption in PCM cell operations is proportional to cell volume to be altered, and in twenty years the advances in litographic techniques had made the nanometer scale a real possibility, bringing PCM power needs in a competitive range. Industrial interest re-exploded, together with the field as a whole.

The need for aggressive scaling comes together with the need for a deeper understanding of the physics of these materials, even more so on the nanoscale, to recognize problems and identify opportunities.

2.1 Structural and phase-change properties

In this section we briefly review the current knowledge of the structure of PCM materials. The most promising and studied alloys for electrical memory applications are those existing on the GeTe–Sb_2Te_3 pseudo-binary tie line in the composition diagram (Figure 2.1). When not specified otherwise, we will focus on the $Ge_2Sb_2Te_5$ (GST) alloy.

As already said, PCM can exist in at least two stable phases, a crystalline and an

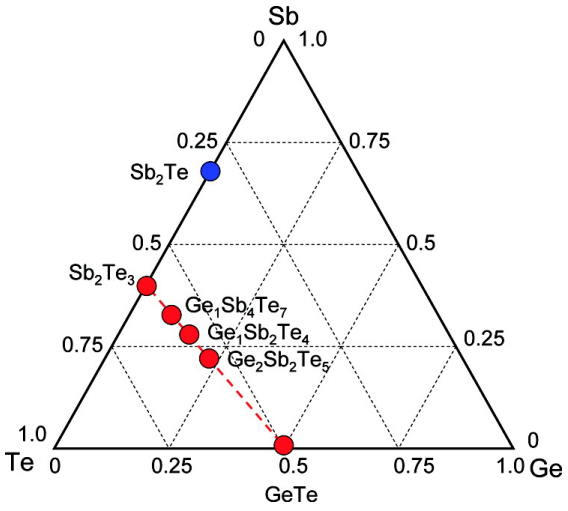

Figure 2.1: Ge-Sb-Te composition diagram, showing the pseudo-binary tie line GeTe–Sb_2Te_3, to which several interesting alloys belong, the promising $Ge_2Sb_2Te_5$ one in particular. Figure taken from [99].

amorphous one, exhibiting remarkably different properties. GST in particular shows three different structural phases, the usual amorphous one together with two crystalline (hexagonal and rock-salt). The rock-salt crystalline is the one involved in the Non-Volatile-Memory (NVM) operations.

A great deal of both experimental and computational effort is devoted nowadays to the study of structural models for PCM, with the comprehension of the amorphous state being the most interesting and challenging issue. In the following, we summarize the current knowledge of the three phases.

2.1.1 Main structural characterization techniques

Optical techniques are among the most used to probe the structures of materials, in both crystalline and amorphous phases, since even the latter can exhibit some short- and medium-range order, despite its lack of long-range periodicity.
A first example of such techniques is the well-known X-Ray Diffraction (XRD), studying the way a beam of x-ray light is scattered by the atoms of a lattice. The angular deviations of the scattered beams are used to acquire information on the material structure at hand, through the examination of the scattering interference.
A related but more sophisticated technique is the X-Ray Absorption Spectroscopy (XAS), based on a composite process in which x-ray photons are absorbed by electrons in the atoms, which results in the emission of photoelectrons that undergo an "internal diffraction" due to atoms in the vicinity of the one of origin. The interference between backscattered and outgoing electronic waves produces energy- and k-vector-sensitive oscillations in the experimental absorption spectra, from which information can be extracted.
In particular, for large values of the k-vector, the XAS photoelectron behavior is dominated by single-scattering processes. This property is exploited in the Extended X-ray Absorption Fine Structure (EXAFS) spectroscopy. The Fourier transform of the spectrum obtained from such techinque, with some extra elaboration, can provide useful information about the atomic Radial Distribution Functions (RDF), in general, and the nearest-neighbor bond lengths, in particular. EXAFS is also a sensitive tool for probing chemical bonding and oxidation state, or for assessing the mean-square relative displacement (MSRD) of the bond lengths.
Somewhat complementarily, when the energy is near the absorption edge, multiple scattering events are more likely, and they are exploited in the so-called X-ray Absorption Near Edge Structure (XANES) spectroscopy. These spectra carry more information about 3- and 4-body correlations in the lattice, thus the coordination number and general surroundings of each atom. They are mostly analyzed together with multiple-scattering simulations, by assuming possible structural arrangements and then comparing the results of the corresponding simulation with the experiments, to find the best fitting model.

2.1.2 Crystalline phase: hexagonal and rock-salt structures

The stable bulk structure of crystalline GST has been identified as a hexagonal lattice, consisting of alternating blocks of GeTe and Sb_2Te_3 [71, 72]. However, GST has also been found to crystallize in a metastable rock-salt structure, which shows several similarities with the atomic arrangement of the material in its amorphous phase. This fact

is sometimes indicated as a possible reason for the short crystallization times of GST, particularly for thin films.
The rock-salt structure is sometimes called also fcc-like, since it displays two intermingled fcc sublattices, as shown in Figure 2.2. The first one is composed entirely of Te atoms,

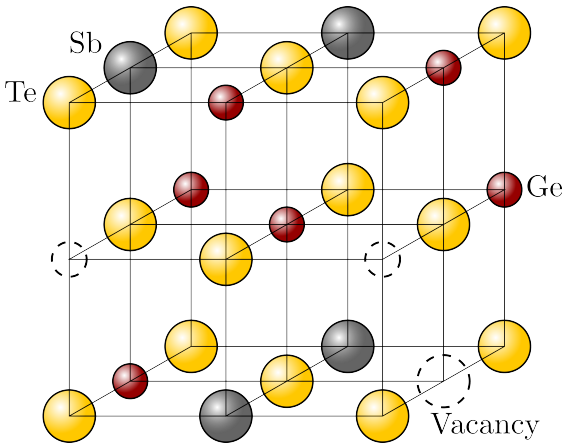

Figure 2.2: Rock-salt structure of crystalline $Ge_2Sb_2Te_5$, as inferred from XRD measurements.

while in the second one a Ge or Sb atom appears randomly at each site, as shown by XRD measurements [6, 9, 90]. Furthermore, each site has a 20% chance of hosting a vacancy, based upon EXAFS investigation [58]. The lattice parameter of this structure is found to be 6.02 Å.
Absorption spectroscopy allowed also to perform a thorough analysis of the chemical bond properties between each pair of atomic species. Figure 2.3 shows the Fourier-transformed data from EXAFS measurements on crystallized and laser-amorphized GST samples, as found by Kolobov *at al.* [58]. The graphs do not directly represent actual real-space RDFs: they still contain an un-corrected phase term related to the photoelectron oscillations under study in the EXAFS spectroscopy, which shifts the peaks to lower positions with respect to the real interatomic distances. Thus, Figure 2.3 simply shows the magnitude of the Fourier Transform of k-space EXAFS data. However, with a proper data elaboration, it was possible to extract plenty of real-space information from these data. We summarize the main results coming from the analysis of such data.
Two different bond types are found, with a Te atom binding with a Ge atom in the first case and with an Sb one in the second. The measured Te-Ge length is 2.83±0.01 Å, while for Te-Sb it is 2.91±0.3 Å.
No Sb-Ge or Ge-Ge bonds were obtained from the spectra, while a second-nearest-neighbor Te-Te peak was found at 4.26 Å. The MSRD for the rock-salt phase is about

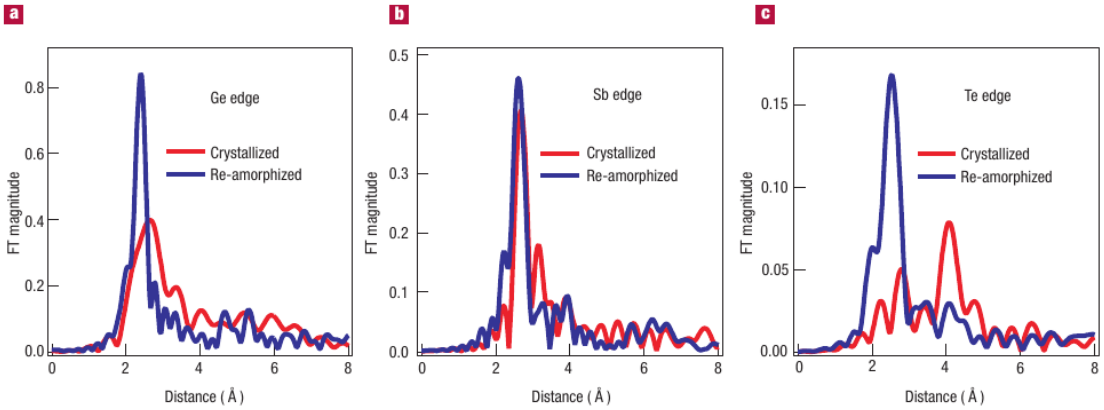

Figure 2.3: Fourier transforms of K-edge EXAFS spectra for Ge, Sb and Te of $Ge_2Sb_2Te_5$ for both the crystalline rock-salt and amorphous states. Figure taken from [58].

0.02 Å^2.

All of these evidences agree with the rock-salt picture depicted above. It is worth mentioning that although one of the two fcc sublattices is generally described as a random distribution of Ge atoms, Te atoms, and vacancies, some *ab initio* calculations suggest that these elements could actually display some clear periodic patterns [118].

2.1.3 Amorphous phase

A sketch of the amorphous structure is shown in Figure 2.4. As reported in [58], the lengths of the Te-Ge and Te-Sb are found to shorten to 2.61 and 2.85 Å, respectively. The second-nearest-neighbor Te peak is still present, although it appears weakened. The MSRD is found to pass from the 0.02 Å^2 of the rock-salt phase to 0.008 Å^2.

These findings appear quite unusual for three-dimensional covalently bonded solids, since typically an increase in disorder corresponds to an increase of bond lengths. However, the amorphous phase of GST is known to show a volume increase with respect to the crystalline one. This fact apparently contradicts the bond shortening evidence, depicting an even more peculiar picture.

Actually, an analogy can be drawn between this behavior and that of molecular solids, which Kolobov *et al.* suggest as an explanation for the experimental evidences above, based on a long-bond-rupture and short-bond-strengthening process. GST can be though of as composed of blocks of atoms, with same-block atoms connected by short, strong bonds and different blocks tied through weaker, long bonds. Upon melting, the weak inter-block bonds break, which makes the short intra-block bonds become even shorter

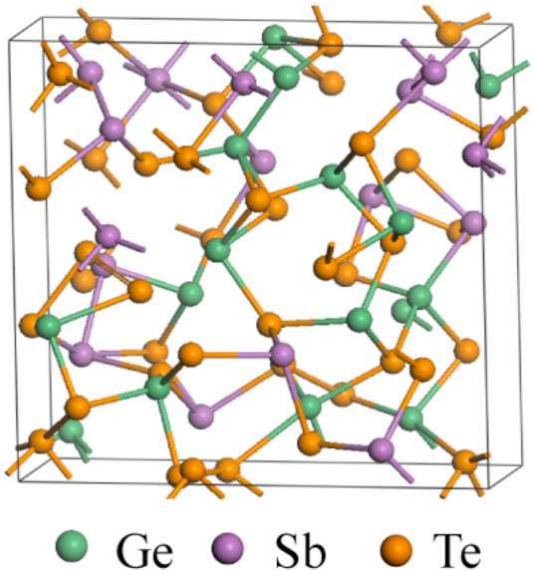

Figure 2.4: Sketch of the amorphous phase of GST from a simulation of the melt-quenching process, including 72 atoms. Figure adapted from [32].

and stronger, implying that the amorphous phase is more *locally* ordered than the rock-salt crystalline one. The rupture of the weaker inter-block bonds allows for a rearrangement of the blocks and can account for the volume increase.
This explanation encompasses the typical bond length shortening, together with the MSRD decrease. Furthermore, it was supported by Raman experiments (see Figure 2.5), where GST Raman modes were seen to become more rigid (i.e. shift toward higher wavenumbers) in the amorphous phase, similarly to the known cases of selenium and tellurium amorphization [30].
To investigate the structural arrangement of atoms on a larger scale than the first nearest-neighbors, the XANES technique was employed, performing simulations of spectra corresponding to several atomic configurations, and comparing them with experiments [58].

The best agreement between the two approaches was found for structures where Ge atoms were allowed to occupy sites of tetrahedral simmetry within the surrounding FCC sub-lattices of Te atoms (see Figure 2.6), instead of the octahedral ones where Ge atoms are found in the crystalline phase. A scheme of this arrangement is reported in Figure 2.7. Such mechanism was proposed by Kolobov *et al.* [58], and nicknamed "umbrella flip model", as a main and characteristic feature of the melting process in GST and as an explanation for the very fast switching properties of the material. However, this topic

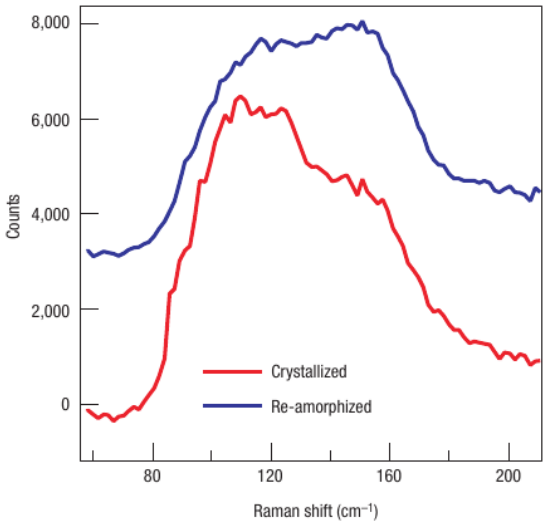

Figure 2.5: Raman results showing amorphous and crystalline GST spectra, the first exhibiting more rigid modes than the second, in agreement with the model proposed by Kolobov *et al.*. Figure taken from [58].

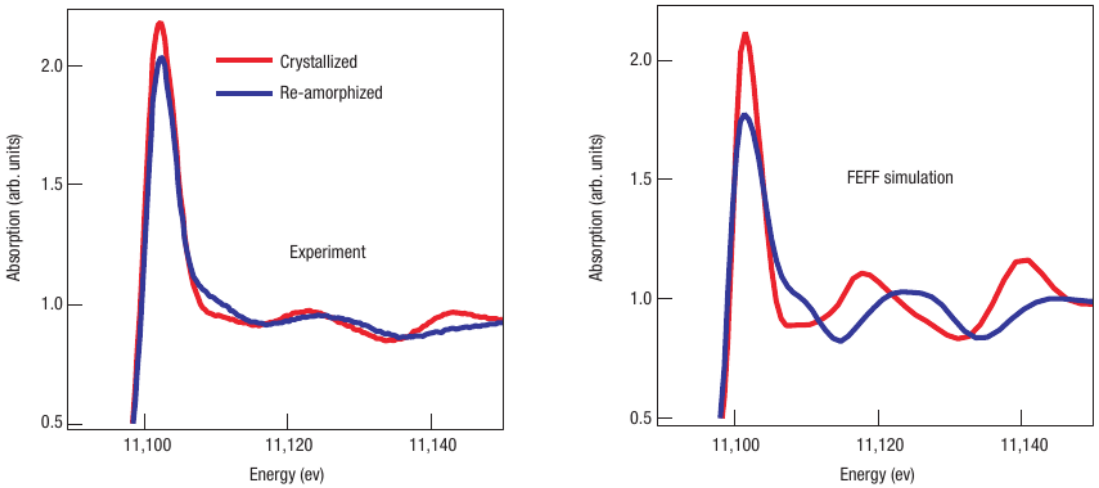

Figure 2.6: Measured (left) and simulated (right) XANES spectra of the Ge K-edge for GST in the crystalline and amorphous phases. Figure taken from [58].

is still quite debated in the scientific community: in fact, several results from *ab initio* simulations in the Density Functional Theory (DFT) approximation generally find that, although 30% of the Ge atoms in the lattice are indeed in tetrahedrally coordinated positions, most of them occupy "defective octahedral" sites, with bond angles close to 90° [28, 10, 119, 42]. This does not cope with the umbrella-flip model described by Kolobov

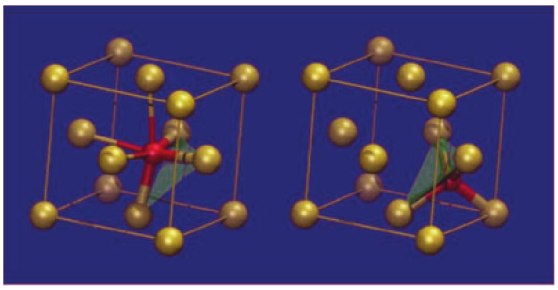

Figure 2.7: Scheme of the octahedral- to tetrahedral-simmetry location change of Ge atoms in GST, as proposed in Kolobov's "umbrella flip" model. Figure taken from [58].

et al..

An alternative picture for the amorphous phase deals with the presence of regular building blocks organized into rings. The crystallization-amorphization process consists of ordering/disordering these blocks, with vacancies playing a helping role in this rearrangement [10].

Another recent model for phase change takes into account the destruction/creation of the so-called resonant bonding between covalently bonded fragments. Specifically, the crystalline state of GST features resonant bonds between orbitals of adjacent molecular units, that are lost in the amorphous phase in favor of purely covalent bonds as the long-range order is lost [69, 115]. Such bonding variation can be the cause of the dramatically different properties of the two phases [44].

As a final remark, it is worth noticing that different amorphous phases (e.g. as resulting from laser re-amorphization, as opposed to thermal evaporation) can produce actually different structures, and comparisons between such cases must be performed carefully [60, 21].

2.1.4 Phase-change kinetics

Proper electrical (or laser) pulses can be used to heat PCM bits and manipulate their states.

In the crystalline-to-amorphous phase transition, enough heat is generated to melt the crystalline sample, leaving it in a liquid state. At this stage, a fundamental role is played by the cooling rate, which determines the structural evolution of the system (see Figure 2.8). If the cooling is slow enough, i.e. it happens in a long time compared to the characteristic times of the crystallization process, the liquid sample can turn into the lower-free-energy crystalline state, with a sharp volume decrease. In the opposite case, the

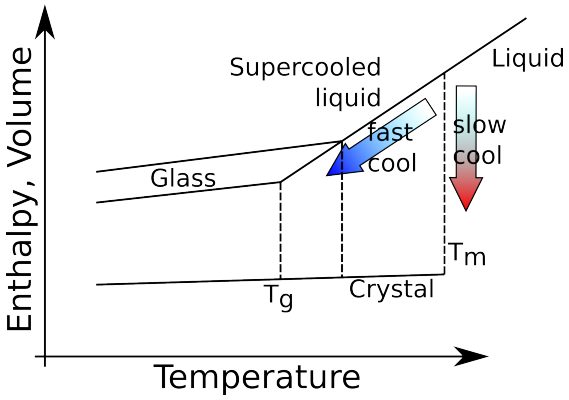

Figure 2.8: Scheme of the glass formation dynamics. See text.

PCM sample becomes supercooled and enters a glassy state around the glass-transition temperature T_g, typically $\sim$ 450 K in GST. Below this temperature, the times needed for further structural rearrangements become very long and experimentally impractical, thus the glass phase forms. In this state, no sharp volume change is seen. Instead, the volume of the final amorphous metastable state depends on the cooling rate, as does its final free energy.

We just mention that more complicate mechanisms for amorphization, that go beyond Joule heating, are possible. In particular, electrical-wind-force-induced amorphization was observed in GST single crystals, as a result of the formation, motion and jamming of dislocations and defects [78].

Crystallization dynamics is determined by the mobility of the atoms in the lattice (which in turn depends upon temperature) and by the difference in free energy between the initial and final phases. In this picture, at the melting point, corresponding to the temperature T_m ($\sim$ 900 K for GST), the liquid and crystalline phases are at equilibrium with each other, which corresponds to a zero free-energy difference. The maximum crystallization rate occurs at temperatures somewhere in between T_g and T_m, to the point that the reduced glass transition temperature $T_{rg} = T_g/T_m$ is sometimes used as an indicator of the nucleation rate.

Typical values of T_{rg} are between 0.4 and 0.84 for generic glass-forming materials, but are usually found in the 0.45-0.55 range for good PCM candidates. This is due to the fact that one of their main desired features for phase-change memories is a fast crystallization ability, in order to obtain fast information writing performances. This property usually coexists with poor glass-forming capabilities: in fact, these materials do not arrange into real random networks of covalent bonds, and instead preserve some

degree of short-range order, as we discussed previously. Furthermore, kinetics studies of crystallization have shown a cooperative effect of electric field and temperature in enhancing the crystallization rate, and determining the nucleation conditions and nuclei shapes [56, 57, 79].

2.2 Electronic properties

2.2.1 Absorption study of the bandgap

Optical spectroscopies of different kinds were used to investigate the bandgap width of GST in all the three phases. The employed technique is based on light absorption by a sample of GST. The intensity of the light transmitted through a slab of material is described by the relation

$$I = I_0 \exp\left(-\alpha \Delta x\right) \tag{2.1}$$

where I_0 is the intensity of the light incident on the slab, Δx is the thickness of the slab and $\alpha = \alpha(\nu)$ the absorption coefficient for the given light frequency ν. The relationship

$$\alpha h\nu = (constant)\left(h\nu - E_g^{opt}\right)^r \tag{2.2}$$

is commonly used to describe optical excitation of electrons from the Valence Band (VB) into the Conduction Band (CB) of a semiconductor. It allows to connect the measurable α to the optical bandgap E_g^{opt}, with h the Planck's constant and r an appropriate constant. Generally, $r = 2$ is assumed for amorphous or indirect-gap crystalline materials, and $r = 1/2$ for direct-gap crystalline ones [121]. Consistently with the indirect-gap case, in agreement with the literature [122], $r = 2$ was used in the cases of GST rock-salt and hexagonal phases (e.g. [67]).
It is possible to plot $(\alpha h\nu)^{1/2}$ as a function of $h\nu$ and extrapolate E_g^{opt} as the intersection with the horizontal axis [76]; this describes the fact that absorption in an ideal semiconductor (i.e. perfect periodic lattice, no defects or impurities) becomes possible only when the energy of the incoming photon is at least equal to the width of the optical bandgap, i.e. the energy difference between the last occupied VB state and the first unoccupied CB state.
Using the method above, E_g^{opt} was determined for the three phases of $Ge_2Sb_2Te_5$ [67, 93, 53], as represented in Figure 2.9 (reported in [100], after a paper by Lee *et al.* [67]). The optical bandgap of the hexagonal and rock-salt phases were both found ~ 0.5 eV, while in the amorphous phase it was found $E_g^{opt} \sim 0.7$ eV. We remind, in passing, that

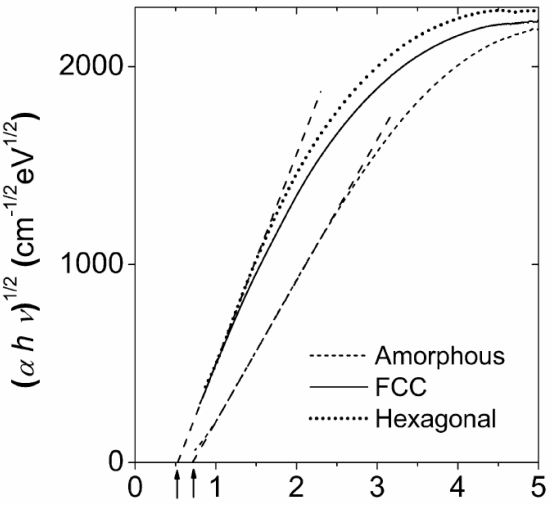

Figure 2.9: Extrapolation of E_g^{opt} from light absorption spectroscopy, see text. Figure taken from [100].

the optical bandgap must not necessarily correspond to the transport bandgap E_g of the material.

It must be noted that Equation (2.2) correctly describes $\alpha = \alpha(\nu)$ only for energies around the bandgap value. Far from that value, the absorption behavior deviates significantly from that relation due to a series of causes, depending on the structure of the material.

First of all, Eq. (2.2) is derived assuming parabolic bands, an approximation which is valid only for electronic transitions from the top states of the VB to the bottom states of the CB. Absorption of high-energy photons defies this assumption and deviates from the behavior described above.

Secondly, a significant absorption of low-energy photons can be seen in the infrared portion of the spectrum, if the material exhibits a high concentration of free carriers. This is indeed the case for the hexagonal and rock-salt crystalline phases of GST, even at room temperature, so that the phenomenon can be interpreted in the approximation of the Drude dispersion model. Since the model includes the conductivity of the material as a fitting parameter, optical investigations can provide an insight into the alternate-current (AC) transport properties of chalcogenide materials, yelding a conductivity parameter equal to 2230 S/cm for the hexagonal phase [74], which is comparable with the bulk measured value of 2400 S/cm from [112]. The AC conductivity of a thin film of the rock-salt phase is much higher than its direct-current (DC) counterpart [67, 38, 70], mostly due to its polycrystalline nature and the subsequent scattering effect of grain boundaries.

Finally, absorption techniques are sensitive to a typical feature of amorphous semiconductors: for these materials the absorption coefficient decays exponentially for decreasing energies in the region below the applicability of the relation of Eq. (2.2), about 0.8 eV for the case of $Ge_2Sb_2Te_5$. This exponential portion of the spectrum (see Figure 2.10) is known as the Urbach edge and is well described by the empirical relation $\alpha = \exp(h\nu/\Delta E)$, with the fitting parameter ΔE found to vary between 50 and 200 meV in GST, depending upon several experimental conditions such as the film growth rate and the oxygen content of the sample. The Urbach edge is correlated with the existence of tails of localized states within the bandgap of the semiconductor, and represents a very important feature of the electronic structure of the amorphous phase for its consequences on charge transport, as will be explained in the subsequent Chapters.

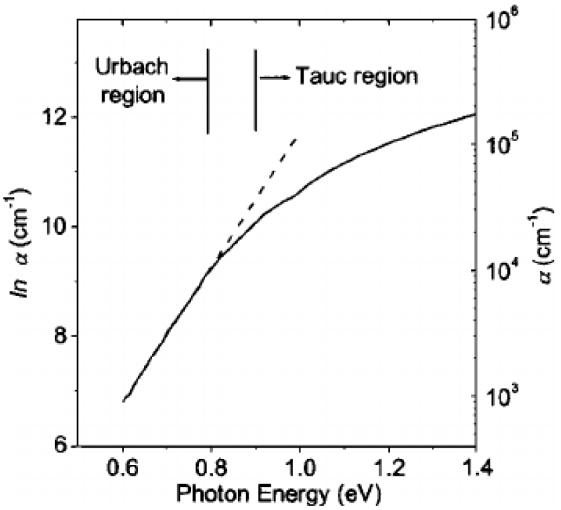

Figure 2.10: Exponential decay of the absorption coefficient α in the Urbach region. Figure taken from [67].

2.2.2 Temperature dependence of conduction properties

In [38], a study of the sheet resistance of GST as a function of temperature highlights the great difference in the electrical behavior of the three phases, particularly the 2-3 order-of-magnitude variation in resistivity between the amorphous and rock-salt samples. This is represented in the left plot of Figure 2.11, where two relevant resistance drops are seen, one at temperature T_1 (amorphous→rock-salt crystalline transition) and one at temperature T_2 (rock-salt crystalline→hexagonal crystalline). It must be noted that

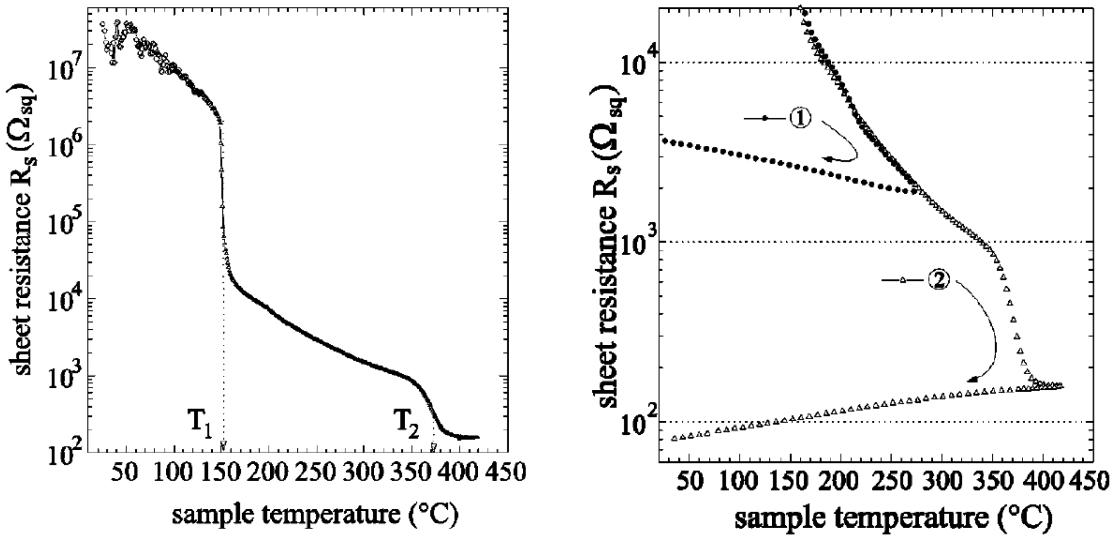

Figure 2.11: Variation of sheet resistivity of an initially 80 nm thick $Ge_2Sb_2Te_5$ film, as a function of temperature. Figures taken from [38].

both plots in Figure 2.11 actually show sheet resistances instead of resistivities. The reason is that the thickness of a GST sample changes accordingly to temperature and phase variations, with the most significant decrease ($\sim 5\%$) in the amorphous→rock-salt crystalline transition at T_1. Thus, the precise thickness was unknown at all points and the resistivity could not be estimated. However, it is clear that a huge resistivity change must occur at T_1, which is a key point of GST exploitation for PCRAM manifacturing.
The right plot in Figure 2.11 shows the effect of cooling in the different phases. In the resistive amorphous phase, the resistance decreases with increasing temperature, due to carrier thermal excitation. If the sample is heated up to a temperature $T \lesssim T_1 \approx 150$ °C and subsequently cooled down, the sheet resistance curve for decreasing temperature is observed to basically correspond to the previous one for increasing temperatures (these curves are not reported in the right panel, since the amorphous region is shown only in the top-left corner of the left panel).
On the contrary, after the amorphous→rock-salt transition, a temperature decrease drives the resistance along a different, slightly increasing curve (marked "1"), while after the rock-salt→hexagonal transition cooling makes the resistance slightly decrease (marked "2"). In both cases resistances are significantly lower than the values measured during the first passage through the same temperature. The two crystalline behaviors are related to mobility variations, and not to changes in the carrier population as it happens in the amorphous phase. In particular, crystal-grain growth plays a role [70]: during heating,

crystal-grain dimensions increase, reducing the scattering due to the presence of grain boundaries. This effect is irreversible, and the sample maintains a lower resistance even during the cooling phase. As regards the different slopes of curves "1" and "2" in Figure 2.11 (right), the reason is often ascribed to the different band properties of the two crystalline phases, as it will be described in the following. In brief, charge transport in the semiconducting rock-salt phase benefits from the thermal excitation of free carriers (and conversely it is weakened by low temperature). On the contrary, the hexagonal phase is a degenerate semiconducting one, so that a temperature decrease does not reduce the amount of free carriers, and instead mobility is enhanced by the reduced phonon scattering.

2.2.3 Concentration, mobility and bandstructure

Hall effect measurements have been employed since the early 1970s to characterize the charge transport properties of chalcogenide materials, together with other features determined through cross-analysis of experimental evidences from other sources (e.g. optical spectra). Conduction in the hexagonal phase was easily determined to be of $p-$type, with a very high concentration of holes, $\sim 3 \times 10^{20}$ cm^{-3} and mobilities at room temperature of ~ 30 cm^2/Vs in thin films [67, 70] and ~ 50 cm^2/Vs in bulk [112]. It should however be pointed out that the mobilities significantly depend on the preparation conditions due to their sensitivity on scattering. Furthermore, low temperature Hall measurements provided a useful insight in the material bandstructure and population statistics: at ~ 5 K, high conductivities were measured, with no carrier freeze-out. The main implication of this is the fact that the Fermi Energy E_F must be located inside the VB, otherwise little or no carriers would be found near 0 K temperature. Thus, hexagonal $Ge_2Sb_2Te_5$ is considered to be a degenerate $p-$type semiconductor [67, 53, 112]. The conduction bandgap E_g is smaller than the optical bandgap $E_g^{opt} \sim 0.5$ eV, since the measured value of E_g^{opt} corresponds to the distance between the in-VB E_F and the CB bottom, which is necessarily larger than the distance between the VB top and CB bottom. This degeneracy condition is possibly related to the highly defective structure of the hexagonal phase, which can induce the formation of a defect band in the vicinity of the VB, overlapping with it and generating the situation just described.
Although not very technologically relevant, the stable and quite well understood hexagonal phase can be used as a reference for the investigation of the other phases, particularly the rock-salt one. With analogous experimental means, the latter was found to share

several features with the former, mainly the $p-$type conduction, a high carrier concentration ($\sim 10^{20}$ cm^{-3}) and a relatively high mobility (~ 1 cm^2/Vs) at room temperature, with the particular values being quite a strong function of the fabrication conditions [67, 70]. The same no-freeze-out behavior of the hexagonal phase was found in the rock-salt phase ($\sim 8 \times 10^{19}$ cm^{-3} at 5 K), suggesting a similar interpretation as a degenerate semiconductor. This last hypothesis, however, is a bit more uncertain, since data from x-ray photoemission spectroscopy experiments in [55] can be interpreted as if the Fermi level of GST-225 was located 0.15 eV above the top of the valence band, inside the bandgap. Another supporting evidence for this alternative picture comes from conductivity measurements as a function of near-room temperature [53], which suggested the existence of an activation energy $E_a \sim 0.15$ eV for transport (i.e., a conductivity $\sigma(T) = \sigma_0 \exp(-E_a/k_B T)$, with k_B the Boltzmann constant, T the absolute temperature and σ_0 a constant value). Nonetheless, according to the non-degenerate interpretation, the standard semiconductor theory would predict hole concentrations in the $10^{16} - 10^{17}$ cm^{-3} range, which should correspond to mobilities of $10^3 - 10^4$ cm^2/Vs for consistence with the measured conductivities, and would represent unreasonably high estimates. It should be finally noted that defect band formation is predicted as a possibility in the rock-salt as well as in the hexagonal phase, due to the many vacancies in the lattice [93].

The amorphous phase is probably the most challenging one. Its transport properties follow a thermally activated behavior, with an activation energy E_a varying between 0.3 and 0.45 eV [67, 53, 66, 132, 14, 48]. If we assume that E_F lies around the middle of the bandgap, then $E_g \sim 0.6 - 0.9$ eV, which is compatible with the measured $E_g^{opt} \sim 0.7$ eV. Furthermore, E_a is found to be a function of both temperature (measurements report a value of 0.41 eV for the range 270-350 K, 0.35 eV for 140-230 K, see [66]) and applied voltage, varying from 0.33 to 0.29 eV for a bias passing from 0.4 to 1.0 V in a PCRAM device [48]. More on the role of E_a and thermally activated processes will be discussed in the next Chapters.

Despite some uncertainties on the carrier sign derived from Hall measurements in resistive amorphous materials (where band conduction is not necessarily the main transport mechanism and the Hall effect can be complicate to interpret), amorphous $Ge_2Sb_2Te_5$ is considered to be $p-$type (according to Seebeck measurements [53, 14]), with a measured Hall mobility of ~ 0.07 cm^2/Vs [14].

Although the presence of localized states in the amorphous material is undisputed, their nature and distribution in the bandgap are less clearly defined, despite the significant

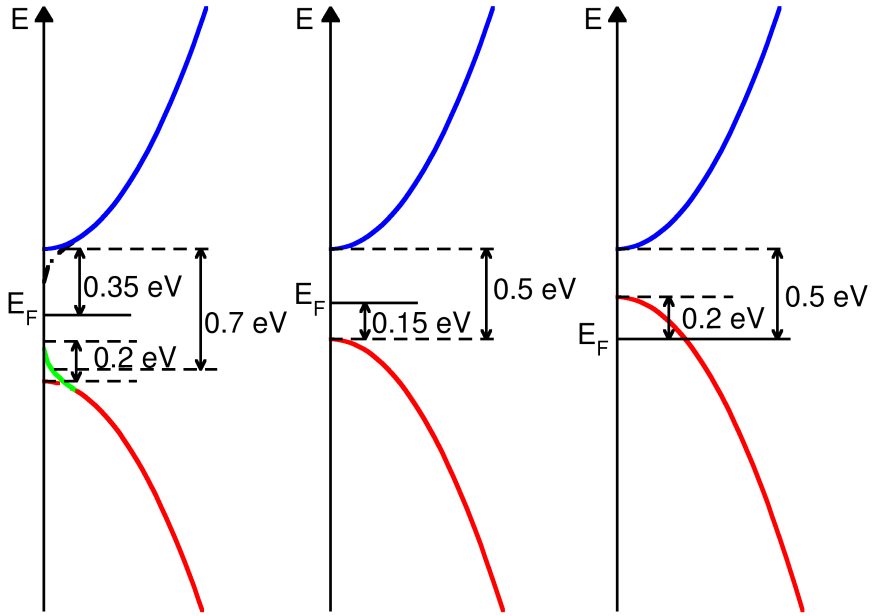

Figure 2.12: Schematic representation of the band structures of the amorphous (left), rock-salt (center) and hexagonal (right) phases of GST, as inferred from considerations about optical spectra and electrical and transport properties.

number of studies which aimed at addressing it [93, 53, 14]. As mentioned in our discussion of the absorption properties by an Urbach edge in the disordered phase, a band tail of localized states most likely exists next to the VB top, which is though to originate from Te lone pairs. Also, Pirovano *et al.* propose that defects in Te-Te chains can produce donor- and acceptor-like levels [93].

A similar model (see e.g. [45] or the review [99] of Raoux *et al.*) uses a slightly different hypothesis and identifies a continuous spectrum of localized states in the whole bandgap, originating from both the VB and the CB and overlapping with decreasing density of states (DoS) near midgap, somewhat as if two Urbach tails were extending from both bands. Figure 2.12 summarizes the band models for the three phases of $Ge_2Sb_2Te_5$.

2.3 Device operations

As previously said, the use of phase-change materials for PCRAMs relies on electrically altering and recognizing the state of a chalcogenide sample, to encode and retrieve data, respectively.
A memory cell can be written (or programmed) employing properly shaped voltage pulses. On a basic level, the most important pulse parameters to control are its amplitude A, its duration t_d and its rising and falling edges, t_r and t_f respectively. At least two different kinds of programming pulses are needed: SET and RESET.

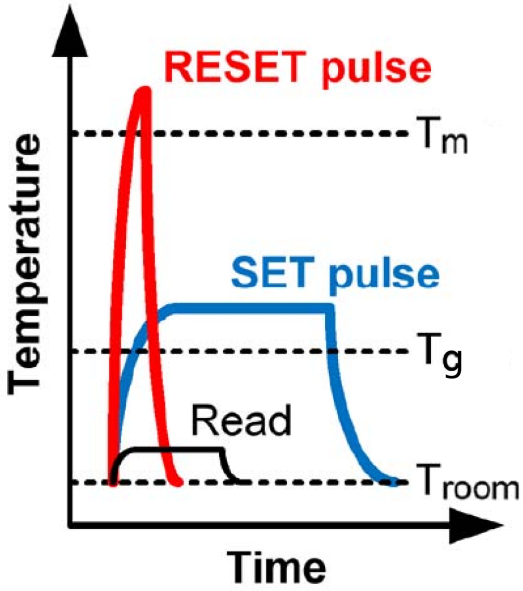

Figure 2.13: Schematic of the temperature profle rises over time, due to SET, RESET and read pulses. Figure adapted from [125].

SET The SET operation is the pulse needed to bring the cell from the amorphous to the crystalline state. It needs a relatively low amplitude, since the purpose is to rise the temperature of the sample above the glass transition temperature $T_g \sim 450$ K, but below the melting temperature $T_m \sim 900$ K (see blue curve in Figure 2.13). Although in some cases a sufficient Joule heating could be achieved for below-threshold voltages, in practice the heating for the SET operation greatly benefits from the Ovonic switching properties of GST, since high current and heating can be generated in the amorphous phase with a bias just above V_{th}.
The SET operation must happen in times long enough for the crystallization to occur,

since a too fast cooling could quench the cell in a disordered state. Thus, t_d and t_f have relatively large values (t_d tens to hundreds of ns, t_f tens of ns). This need makes the SET operation the technological "time bottleneck" of PCM devices, i.e. the most time-consuming among those needed for full control of PCRAMs.

RESET The RESET operation induces the transition from the crystalline to the amorphous phase. Since this requires to melt and rapidly cool the memory cell (red curve in Figure 2.13), the RESET voltage pulse has a higher amplitude and shorter duration and edge times than the SET pulse. The cell can easily be heated above the temperature T_m, and times of the order of $t_d \sim 50$ ns and $t_d \sim 2$ ns allow to suddenly quench the liquid GST into a disordered phase. While being a very fast step in the memory operations, the RESET pulse is a very power consuming one.

Read Reading is usually achieved with quite short, low-amplitude voltage pulses. The main requirement is that, during read operations, there must be no risk to alter the state of the memory bit; in general the reading voltage is a fraction of V_{th}, such that the read-out happens in the region where the resistance difference between SET and RESET state curves is the largest (see e.g. the scheme in Figure 1.1). Other important features for a proper READ operation are low power consumption, which is in general correlated with a low pulse amplitude, and a high-signal-to-noise ratio, which is usually better for higher READ voltages. The trade-off between all of these trends is an important point in memory device design.

The particular numbers of the figures of merit described above depend significantly on material and design choices. However, the trends usually remain the same, i.e. slow, low-power SET versus fast, power-consuming RESET. A great effort is devoted to both optimizing material choice and to explore the scaling potential of chalcogenide materials [130, 129].

2.4 Summary

This Chapter was a brief overview of the basic physics of phase-change materials. The current understanding of the structural properties of the crystalline and amorphous phases was reviewed, in order to gain insight into the mechanisms of phase-change between the amorphous and the rock-salt structures, which is at the basis of device operations.

The electronic properties of GST in all phases were then reported, as resulting from both optical and electrical measurements. These investigations unveiled important features of GST such as the presence of localized states in the amorphous phase, the bandgap width, and the basic macroscopic transport properties of all three phases, as well as their behavior as a function of temperature. These information will come into hand in the following Chapters.
Finally, the memory operations for PCM-based devices were described.

Chapter 3

Conduction in amorphous phase-change materials

Electrically controlled PCM memories have already been successfully brought to industrial production [1]. However, the physics underlying the conduction properties of the amorphous phase is not fully understood, particularly as regards the Ovonic switching phenomenon, which at present is still a debated issue (for a review, see e.g. [21, 80]). Thus, the understanding of the transport physics in amorphous chalcogenides is very important, both at a fundamental level, and as a means to devise better modeling tools for design and to gain insight into new possible device optimization opportunities.
In this Chapter we will review the main experimental facts carrying useful information about the nature of charge transport in GST, and the present theories for the modeling of those evidences, with focus on the amorphous phase and its switching properties.

3.1 Electrical conduction and switching evidences

Chalcogenide glasses exhibit a rich variety of phenomena that a complete model should account for. In this section we review the experimental evidences related to transport in chalcogenides that can help in shedding light on the underlying physics.
First of all, chalcogenides present a typical shape of the IV curve (see Figure 3.1).

At low applied voltages the curve is linear (green region in Figure 3.1); for higher voltages, instead, the device responds with a current of the exponential form $I \propto \exp(cf(V))$, with the coefficient c often called Sub-Threshold Slope (STS) and $f(V)$ a function of the bias, generally proportional to V or $\sqrt{V}$ (Figure 3.1, orange region).

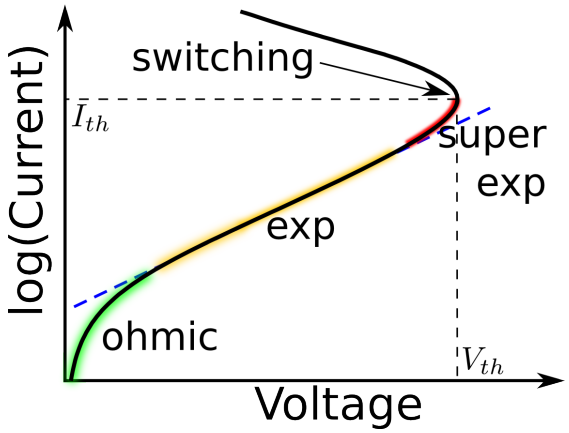

Figure 3.1: Scheme for the typical IV characteristic curve of a PCM device in the amorphous phase. The linear (green), exponential (orange) and super-exponential (red) conduction regions are highlighted. The blue dashed line represents an exponential behavior.

The exponential regime can provide a lot of information about the mechanisms ruling conduction, as experimentally investigated by Ielmini and Zhang in GST industrial memory cells [48]. Measurements were performed both as a function of the applied bias V for several temperatures T and, complementarily, recording the current as a function of the inverse thermal energy $1/k_BT$ (k_B is the Boltzmann constant), in order to draw the so called Arrhenius plot. This was repeated for several applied biases. The results are shown in Figure 3.2. Varying the temperature between 25 and 85 °C, a great current increase

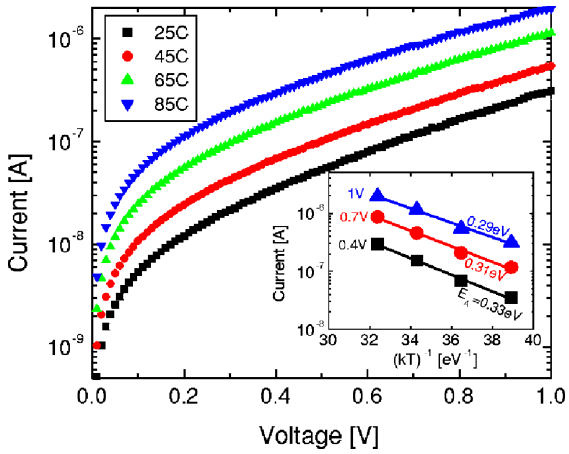

Figure 3.2: Study of sub-threshold conduction in a GST memory cell device. The large picture shows IV characteristic curves measured for four different temperatures. The inset shows the Arrhenius plot for the same device, at three fixed values of applied voltage. Figure taken from [48].

was observed, supporting the idea that conduction in chalcogenide glasses is a thermally activated process, characterized by a relatively large activation energy E_a [63]. Furthermore, the Arrhenius curves in the inset of Figure 3.2 were used to extract the value of E_a, assuming the functional form $I = I_0 \exp(-E_a/k_B T)$, that describes thermally activated processes. The fit provided $E_a = 0.33, 0.31, 0.29$ eV for bias $V = 0.4$,0.7,1.0 V respectively, suggesting that the applied electric field lowers the "activation barrier" that rules the dynamics of the conduction mechanism.
As proposed by Ielmini and Zhang, this behavior is consistent with a trap-limited transport picture in which electrons from localized trap states in the bandgap contribute to conduction after being thermally excited to the conduction band, and are later recaptured by a different trap. Representing the trapping potentials of the starting and final traps as Coulombic ones, the effect of the electric field is to lower the barrier between them, thus enhancing the efficiency of the thermal emission (as in the Poole effect, see section 3.2). The relationship

$$I \propto \exp\left(-\frac{E_a - qV\Delta z/(2u_a)}{k_B T}\right) \tag{3.1}$$

provides an empirical description of the phenomenon in the exponential portion of the IV curve, with q the absolute value of the electron charge, $\Delta z/2$ half of the typical inter-trap distance (assumed to be the position of the inter-trap barrier maximum, located at the midpoint) and u_a the thickness of the amorphous sample, so that $F = -V/u_a$ is the average field in the device.

At high voltages the current further increases with respect to the exponential regime, and the IV curve becomes super-exponential (red region, Figure 3.1). This region is followed by the Ovonic Threshold Switching (OTS) phenomenon, occurring at a threshold voltage V_{th}, and the IV characteristic exhibits an S-shaped Negative Differential Conductance (S-NDC) [111]. When a high enough voltage $V > V_{th}$ is applied, the resistance of the device suddenly drops, reaching a conductivity comparable to that of the crystalline phase, and in phase-change materials crystallization usually follows very closely. Typical values of V_{th} are in the 1-2 V range for industrial cells, sometimes higher for novel prototypical devices. However, it was also pointed out that, although V_{th} is the operatively relevant parameter, the material is actually characterized by an approximately geometry-independent threshold electric field F_{th}, typically found in the range $\sim$ 50–100 V/μm in [61, 130, 129].

It is important to remark that the full smooth S-shaped NDC in GST can only be indirectly measured or theoretically predicted. In fact, the experimental setups for IV measurements consist of both the resistor R_{PCM} representing the device and a series resistance R_s, given by a combination of several possible contributions, most likely a load resistance used to limit the current after the switching, and/or parasitic resistances present in the circuit. When the OTS occurs, the system switches from the so-called "off-state" in the exponential sub-threshold regime, to the "on-state" belonging to the upper branch of the IV S curve, as determined also by the series resistance R_s. At this point the material is still amorphous.
Another switching-related phenomenon is the existence of a holding voltage $V_h < V_{th}$, which can be measured for slowly crystallizing materials, so that the onset of the phase change does not prevent its observation. It corresponds to the minimum bias that still retains the on-state in the device after an OTS event, i.e. if the on-state amorphous device bias is reduced to a voltage V such that $V_h < V < V_{th}$, the on-state is not lost. The off-state is recovered only if a voltage lower than V_h is applied, with the system moving back to the low-conductivity branch along a load line parallel to the one followed during the switching event.
The above properties were described from a steady-state perspective. If studied in their dynamics over time, these effects are associated to additional transient behaviors. The OTS is characterized by a delay time t_{delay}: if a bias above V_{th} is applied to the device, the on-state does not follow immediately, and requires instead a time t_{delay} to occur [46]. Furthermore, the delay time decreases exponentially for increasing over-voltage $V_o = V - V_{th}$ [65].
The holding phenomenon shows a time dependence, too. First of all, the conductivity decreases in a characteristic time t_h. Secondly, the on-state is lost in a characteristic off-state recovery time $t_{recovery}$, in the sense that after removing the bias from the on-state, if a new bias V is applied within a time $t_{recovery}$, it only needs to be $V > V_h$ to reach the on-state again. On the contrary, a voltage $V > V_{th}$ is needed once again for times longer than $t_{recovery}$ [96, 7, 46].

3.2 Review of theoretical models

This section introduces the main theoretical models describing charge transport in amorphous chalcogenides. Their physical interpretations for the shape of the IV and the switching phenomenon are presented. The merits and limits of these models are briefly

discussed.

3.2.1 Double injection

In 1969, Mott [75] and Henisch [43] proposed two similar models for the switching mechanism in amorphous chalcogenides. Based on the high concentration of free carriers ($10^{18} - 10^{19}$ cm^{-3}) found in the on-state of switching glasses [7], they discarded a bulk origin (thermal excitation, impact ionization) for such generation, and assumed a key role of injection from the contacts.
In the so-called double injection model, holes and electrons coming from the contacts recombine in the material within deep traps. This at first generates thin space-charge layers near the contacts, negative (positive) on the electron (hole) injection side. Increasing the applied electric field, the charged region extend far enough from the contacts to overlap in the middle of the device, creating a neutral region. The field near the contacts increases and so does injection, triggering the switching event. The on state corresponds to the neutralization of most of the device, with the exception of very thin potential barriers at the contacts, which however can easily be tunneled through.
This model captures the holding voltage V_h dependence on the film thickness, and quite correctly predicts that $V_h \approx E_g/q$, with E_g the electronic bandgap. Some drawbacks of the model are instead the unclear mechanism of barrier formation and the origin of the rapid carrier recombination in the bulk, particularly as the bulk actually reduces to a nanometric thin-film.

3.2.2 Generation and Recombination

In 1980, Adler *et al.* proposed a model analyzing the generation-recombination dynamics of carriers in a material hosting at the same time defects of three different species (positively or negatively charged, and neutral) [8], according to the Valence Alternation Pair (VAP) model (see e.g. the review in [21]).
They formulated a set of kinetic equations: they wrote the continuity equations for both electrons and holes, including capture by charged defects (neglecting that from the neutral ones, assumed much less effective), a monotonically field-enhanced generation mechanism $G(F)$, proportional to the number of carriers, and a thermal generation rate G_{th}. They also wrote rate equations for the defect populations of the three species, accounting for the interactions among themselves (chemical change) and with the carriers (trapping and de-trapping).
Although the theory is in principle symmetric with respect to the majority carrier sign,

the numerical coefficients were chosen so that the hole concentration p was much greater than the electron concentration n. Assuming an impact-ionization phenomenon, modeled with a generation rate $G(F) = pg(F)$ (with $g(F)$ the monotone function of the electric field), and considering only a hole population and an equally relevant effectiveness of charged traps, it was possible to find the simplified steady-state homogeneous solution

$$p = \frac{G_{th}\tau_p}{1 - g(F)\tau_p} \tag{3.2}$$

for the hole concentration, with $\tau_p = 2/(\beta_p N_{tot})$, β_p and N_{tot} being the capture coefficient for holes by the positively-charged defects and the total number of defects, respectively. The current density was calculated according to $j = qp\mu_p F$, with μ_p a given hole mobility value. The threshold switching is identified by the condition $g(F_{th})\tau_p = 1$, with F_{th} the threshold field, corresponding to an infinite concentration of holes in the material, as in an avalanche breakdown. Also, the linear, exponential and super-exponential regions of the current-voltage characteristic curve are correctly described. Reference [8] reports the more complete and complex solution for the whole set of variables of the model (electrons and defect species), without the above approximations.

Eventually, the model led to the prediction of μm-wide conducting filaments, a condition that is not compatible with the cross-sections of modern devices [8].

As will be discussed later, the first historical attempts to explain the switching behavior invoked a thermal origin. Adler's approach was the first one to show that threshold switching could be explained in terms of a purely electronic phenomenon, without involving a thermal component, a condition which is quite widely (but not totally) accepted nowadays.

A similar model was employed later by Pirovano *et al.* [93], in the context of standard simulations of semiconductor devices with the drift-diffusion model. The continuity equations for carriers of both signs were coupled with the Poisson and the heat-transport equations. The population dynamics of different defect species were not studied in this case, but impact ionization was once again assumed as the mechanism for carrier generation, since it features the desired properties of monotonic direct dependence on the electric field and proportionality to the concentration of carriers. Recombination via trap states in the bandgap was included. The model predicted the switching point with a good agreement with experimental data, even more so after the inclusion of the temperature analysis in the device and the implementation of its effect on the impact ionization process. This approach also allowed to predict the on-state branch of the IV curve. Furthermore, the full coupling with 3D standard semiconductor equations allowed for

the integration of a crystallization algorithm, able to predict partial crystallization and percolative effects [101].
Despite the good predictive power of the generation-recombination approach in all the versions above, its physical correctness is questioned, mainly due to the choice of impact ionization as a triggering phenomenon: the carrier mean free path is believed to be way too short in defective chalcogenide glasses to allow carriers to travel freely long enough to gain the energy for the switching-enabling generation processes, since a quantity at least equal to $E_g/2 \sim 0.3$ eV would be needed.

3.2.3 Poole/Poole-Frenkel conduction

The generation-recombination approach employed a set of balance equations to describe how the carriers moved among energy levels, having different roles on conduction. An alternative approach is to directly assume an energy distribution, such as a Fermi-Dirac function (or more simply a Boltzmann function, in the appropriate conditions) for the carriers, which accounts for the thermal excitation to higher-energy levels.
In the works by Ielmini *et al.* [47, 45], this hypothesis is made, together with the assumption of a Poole conduction regime for defective semiconductors. Carriers are hosted by localized states in the bandgap, described by Coulombic potential profiles. Carriers can contribute to the current flow when they are excited above a given energy barrier, separating two traps, with its top located at energy E_c: the effective barrier felt by a carrier occupying a trap state at energy E_t is $E_a = E_c - E_t$. Consequently, in the absence of an electric field, the number of carriers succeding in overcoming the activation barrier is described by $n \propto \exp\left(-E_a/(k_B T)\right)$. The applied electric field comes into play by lowering the barrier by a field-dependent quantity $\Delta E_a(F)$, so that charge emission is enhanced, according to $n \propto \exp\left(-(E_a - \Delta E_a(F))/(k_B T)\right)$.

With respect to the field effect, two different regimes are classically identified (see Figure 3.3): in the so called Poole-Frenkel conduction, the density of defects in the material is low, and the tails of the Coulombic potential do not overlap significantly, as represented in the left panel of Figure 3.3. This condition mathematically results in a square-root dependence of the barrier lowering upon the electric field F, $\Delta E_a^{PF} = \gamma\sqrt{F}$, with $\gamma = q\sqrt{q/(\pi\epsilon)}$ and ϵ the dielectric constant of the material. In the high trap density case, charge transport is described by Poole conduction: the traps are close to each other, and the Coulomb overlap contributes to the barrier lowering, finally yielding a linear F-dependence of the form $\Delta E_a^P = qF\Delta z/2$ (Figure 3.3, right panel). The coefficient proportional to $\Delta z/2$ implies that the maximum of the barrier is assumed to be at

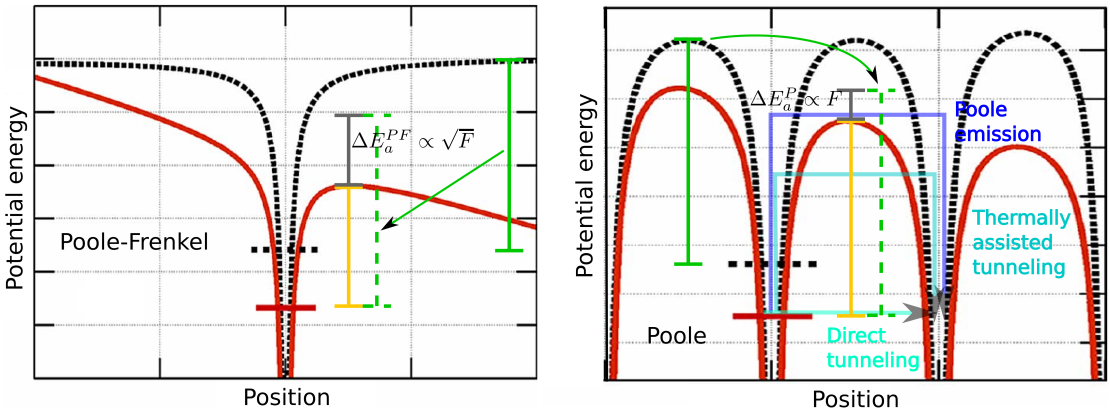

Figure 3.3: Schematic representation of the Poole-Frenkel (left) and Poole (right) regimes. An electric field linearly lowers the potential profile due to Coulombic traps. The barrier seen by a carrier trapped in initial site (marked with the horizontal line) is decreased by a quantity indicated by the grey segment, given by the difference of the barrier value without (green) and with (yellow) the field. Figures adapted from [26].

halfway between the two traps.

The Poole condition is quite convenient in describing the subthreshold activated behavior observed in [48], as explained in Section 3.1.

Ielmini and Zhang [47] proposed a model based on the Poole picture, in which emitted carriers travel until they reach the next trap, located a tipical distance Δz apart, and are immediately re-captured (a slight difference from the original Poole conduction mechanism, in which the carriers were in general allowed to travel according to their own mean-free-path). The model was originally derived for electrons, but it can easily be adapted for holes, and one or both populations can be used for the overall transport description.

The time needed for carrier emission was calculated for both the directions along (forward, $\rightarrow$) and against (backward, $\leftarrow$) the electrostatic force, since in the two cases the electrons need to overcome different activation barriers ($E_a - \Delta E_a^P$ and $E_a + \Delta E_a^P$, respectively, see the Poole panel in Figure 3.3). The calculation yields

$$\tau_{\rightleftarrows} = \tau_0 \exp\left(\frac{(E_c - E_t) \mp q(\Delta z/2)(V_a/u_a)}{k_B T}\right) \tag{3.3}$$

where the right-pointing (left-pointing) arrow corresponds to the minus (plus) sign, τ_0 is a characteristic time, usually taken in the range of phononic frequencies (10^{13}-10^{14} s^{-1}), V_a is the bias applied to the device in the amorphous phase and u_a is its thickness so that, assuming a constant electric field, $F = -V_a/u_a$.

By calculating the average electron velocity in both directions as $v_{\rightleftarrows} = \Delta z/\tau_{\rightleftarrows}$, it was possible to find the current contributions $dI_{\rightleftarrows}$ due to carriers starting from an energy level in an infinitesimal interval dE_t around E_t, namely

$$dI_{\rightleftarrows} = qAn_t\left(E_t\right)\frac{\Delta z}{\tau_{\rightleftarrows}}dE_t = qAn_t\left(E_t\right)\frac{\Delta z}{\tau_0}\exp\left(-\frac{\left(E_c - E_t\right) \mp q(\Delta z/2)(V_a/u_a)}{k_B T}\right)dE_t \tag{3.4}$$

where A is the cross-sectional area of the device and $n_t(E_t)dE_t$ is the concentration of electrons in trap levels with energy between E_t and $E_t + dE_t$. Then, the substitution $n_t(E_t) = N_t(E_t)\exp\left(-(E_t - E_F)/k_B T\right)$ was performed, with E_F the Fermi level and $N_t(E_t)$ the density of trap states per volume, coming from the approximation of the Fermi-Dirac distribution function as a Maxwell-Boltzmann one. Finally, integrating Eq. (3.4) it was found

$$I = qAN_{t,tot}\frac{\Delta z}{\tau_0}\exp\left(-\frac{E_c - E_F}{k_B T}\right)\sinh\left(\frac{q\Delta z V_a}{2u_a k_B T}\right) \tag{3.5}$$

where $N_{t,tot}$ is the total trap concentration in the material.

The model shows a remarkable agreement with subthreshold experimental data. The

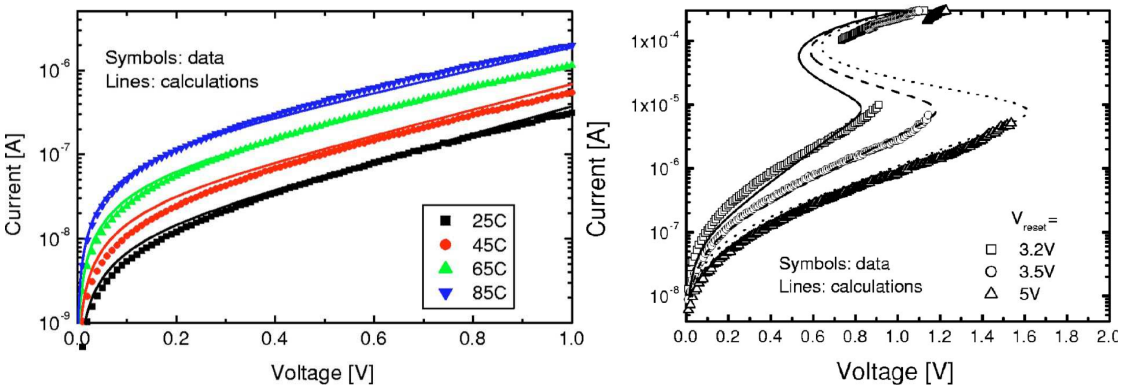

Figure 3.4: Comparison between model calculations and experimental data, for different fixed values of device temperature (left panel) and different device amorphization thicknesses u_a, due to increasing amplitudes of the RESET pulse (right). Figures taken from [47].

left panel of Figure 3.4 shows and well compares theory and data from a device kept at four different equally spaced values of temperature, from 25 to 85 °C, recovering the expected activated behavior as in [48]. In the right panel, the model is fitted to data from the same cell after RESET pulses with three increasing amplitudes V_{reset} =3.2, 3.5, 5 V, likely corresponding to three consequently different amorphization thicknesses u_a. The fit was generally good in the subthreshold region, and provided u_a =18, 27, 39 nm,

respectively for the three cases. In particular, the linear region is well described by the model thanks to its hyperbolic sine form, coming from the difference of the de-trapped carrier flowing in the two directions.

Although in the cases shown here the Poole assumption provided good results for the subthreshold data of several industrial devices, some experimental IV curves exhibit a more $\sqrt{F}$-like behavior. In some cases the two regimes coexist, F for low fields, $\sqrt{F}$ for high fields, and more generally it is not always easy to ascertain the dominance of one or the other. Calderoni *et al.* and Betti Beneventi *et al.* addressed this issue by formulating a field dependence for ΔE_a, able to make the transition between the two conditions [26], or by considering the full 3D dependence of the Poole-Frenkel and predicting both behaviors as a result [19].

The theory up to this point, however, cannot predict any switching. Therefore, Ielmini and Zhang also assumed the existence of two populations, occupying deep and shallow trap level, and a Fowler-Nordheim tunneling mechanism coupling the two levels. At the switching point, the device is divided in two regions along its length, an "OFF" region near the injection contact and an "ON" region near the exit contact, the first exhibiting equilibrium population and electrons mostly in the deep levels, the latter in out-of-equilibrium conditions, with many electrons excited to the shallow states. This is possible because at high fields the tunneling mechanism injects carriers from the electrode and the OFF region into the ON region, creating the off-equilibrium situation that allows a high current with a low bias, and thus the snap-back. In the right panel of Figure 3.4, the simulations reproduce the switching point (i.e. the discontinuity of the experimental data) with reasonable accuracy.

Later, Ielmini proposed a somewhat different mechanism enabling the off-equilibrium regime [45]. The localized states are assumed to be continuously distributed in the bandgap (not only on two trap levels), populated according to a Maxwell-Boltzmann distribution. The integration leads to a constitutive IV equation basically equivalent to (3.5), although it is described as involving an elastic direct tunneling and a thermally assisted tunneling, together with the Poole emission mechanism. The main difference is the way higher-energy levels become populated: an equation for carrier energy balance is written for each infinitesimal thickness dz in the film, accounting for carrier energy gain by the electric field and relaxation due to phonons. The amount of energy gain of the population is represented by a variable pseudo Fermi level $E_F(z)$ along the device.

A steady-state solution is found which predicts the OTS and finds a position-dependent electric field, with a high field in the first part of the device that gradually decreases and finally reaches a low constant value in the second part, consistently with the (more

schematic) OFF/ON regions of the previous model.
The Poole picture described above provides good agreement with experiments, with respect to subthreshold/switching features, as well as with other phenomena which can be interpreted in the same framework, such as the transient behavior of the threshold switching [65] and the resistance drift [31, 33]. Although the physical assumptions seem very reasonable, some can be questioned, such as the role of the Δz parameter, to which quite high values are attributed (more than 5 nm), or the use of the pseudo Fermi level $E_F(z)$ as measure of the average energy. Furthermore, the model neglects the self-consistency between the Poisson equation and the local population, in such a way that a localized charge must exist in order to justify the calculated values of a varying field [47] or of the pseudo-Fermi level [45].

3.2.4 Field-induced nucleation

An alternative, very different approach to the switching problem was proposed by Karpov *et al.* [52]. They started from the observation that the conductivity in the after-switching regime and in the crystalline phase were very similar. This suggested that the switching itself could be due to crystallization, favored by a high field.
In this picture, the switching begins with the nucleation of a long crystalline cylinder, or "needle", with a high aspect ratio h/R. This concentrates the field lines in the vicinity of its tip, in a lightning-rod fashion, according to $F_{corrected} = F/n$, with n a correction factor dependent on the nucleus shape. The way this nucleus enlarges is studied thermodinamically through the free energy

$$G = S\sigma - V\mu + W_F \tag{3.6}$$

to find the favored shape of the new crystalline embryos, which contribute to lengthen the filament. Here S and V are the surface and volume of the new crystalline portion, σ and μ the surface energy and the change in chemical potential, respectively. The last term is the electric field energy $W_F \propto VF^2$.
By maximizing the free energy, it is possible to find the nucleation barrier that separates the region of embryo sizes which will tend to increase from those that will tend to decrease. It is found that the filament lengthens by a series of crystalline spheres at low fields [52, 51], since they feature the lowest nucleation barrier (see Figure 3.5, left). At high fields, instead, new needle embryos are preferred [51]. This is due to two main reasons: first, the strong symmetry-breaking effect of the field lowering the nucleation barrier and second, the increased number of geometric degrees of freedom the needle shape has with

respect to the sphere shape (2D $h - R$-space versus 1D R-space); this latter condition helps the needle case in finding a lower-free-energy evolution path.

The switching occurs when the filament shunts the two electrodes. The model provides

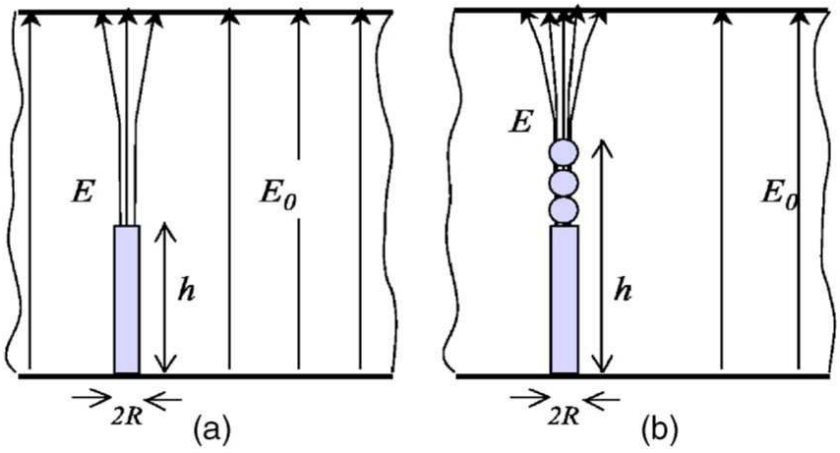

Figure 3.5: Formation of a crystalline filament by nucleation of an initial cylindrical embryo (a), followed by a chain of spherical crystallites (b), which is favored for relatively low electric fields. For higher fields, the simmetry-breaking effect makes the cylindrical-needle shaped more convenient as regards the free energy. Figure taken by [52].

a relation between V_{th} and the switching time τ_s, and a formula for $V_{th} = V_{th}(l, T)$ is found, with l the device length and T its temperature, that provides good agreement with experimental data. Furthermore, the model predicts the possibility of switching for applied biases below V_{th}, if kept for a long enough time, which well compares with dedicated experiments at several temperatures [51].

The stability of the crystallized region is investigated after the removal of the electric field, as shown in Figure 3.6 (left picture), where the two curves refer to the free-energy of the system with (on) and without (off) an applied field. The analysis starts in an applied field condition, from a stable crystalline nucleus i.e. one such that the field-on free energy corresponding to its dimension decreases if its size grows. Once the field is removed, the system moves to the field-off curve, at the same crystal size, and the slope of such new free-energy curve determines the fate (vanishing or growing) of the crystalline nucleus [52, 51, 79]; e.g. a nucleus having dimension X such that $X_E < X < X_0$ (with X_E and X_0 the sizes of maximum field-on and field-off free-energy, respectively) under an electric field will tend to grow (on curve), while upon removal of the field it will start to shrink (off curve).

A similar analysis is applied to purely OTS materials, i.e. those able to switch but unable to crystallize subsequently to the switching [79]. The minus sign in front of

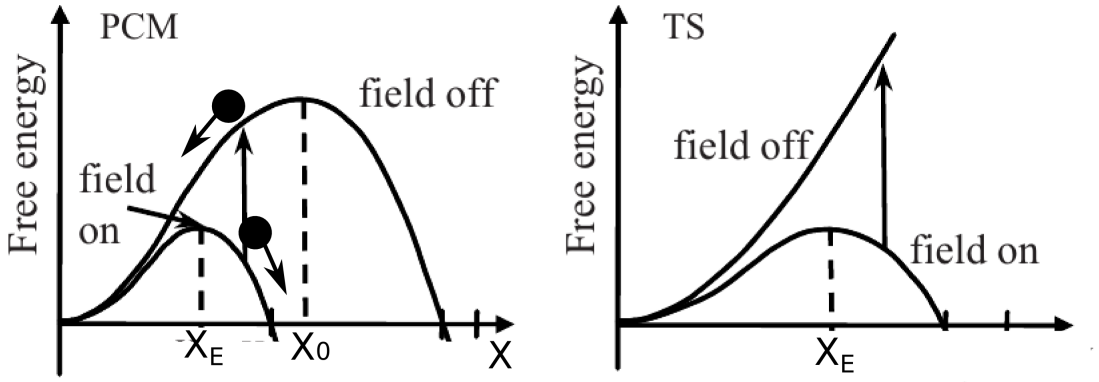

Figure 3.6: Free energy graph for a PCM (a) and OTS (b) device. See text. Figure adapted from [79].

the volume contribution to G (Equation (3.6)) is substituted with a plus, based on the observed higher chemical stability of the amorphous phase with respect to crystal in OTS materials. As a consequence, there is no filament size providing stability after turning the field off (see Figure 3.6, right panel), but it is possible to find the minimum holding field F_h for which a conductive filament still exist, namely the minimum F_h for which the difference between the free energy of the system with and without the filament is $\Delta G \leq 0$. This is done including a screening length λ within which the field is weakened around the filament, when it shunts the electrodes.
The model succesfully describes several experimental switching-related evidences, both as a function of device geometry or temperature and as a function of time (e.g. oscillating behavior [79]). However, its correctness is more uncertain as regards the charge transport behavior, since no explicit prediction of a complete IV was presented by the authors.

3.2.5 Small polaron hopping

Emin [36] proposed a model in which charge transport was ascribed to small-polarons, i.e. carriers that strongly self-localize, due to the polarization they induce distorting the surrounding lattice [13, 14]. The model was current-driven and represented a device as a 1D series of sites, with the central ones modeling the amorphous chalcogenide and a group of left- and right-most sites representing high-mobility contacts.
Three different transport timescales are involved, due to the characteristic transition rate constants: R for the hopping between central sites, R_f for the quasi-free motion between sites of the same contact, and r representing the transition between the quasi-free and small-polaron states at the chalcogenide-contact boundaries. These are such that the

inequalities $R_f \gg R \gg r$ hold true.
Carriers must go from one contact to the other, passing through the whole chalcogenide chain. The main bottleneck is the low conversion rate r at the contact-chalcogenide interfaces. As higher currents are pushed into the device, carriers accumulate near the exit contact, forming a peak in the charge density. The small-polaron self-trapping phenomenon is unstable under high charge-density conditions, since closely-located charges tend to mutually balance their respective effects on the lattice, so that the polaron confinement effect is weakened. Thus, the highly charged end of the chalcogenide region can be turned into a contact-like high-mobility region. The remaining amorphous chain is nonetheless unstable itself, so that this process can trigger the threshold switching in the whole domain, once a sufficient current is injected in the device.

3.2.6 Thermal component

As we have shown, the lattice temperature plays an important enabling role in the conduction of amorphous chalcogenides, due to the activated nature of the process. Therefore, in the early years of study about switching glasses, it seemed reasonable that the heating of the atomic structure was at the basis of the switching event. In many works, a purely thermal origin was assumed for the switching, often in the hypothesis of a conductivity exponentially dependent on $-1/T$.
The phenomenon was described as a breakdown mechanism in many variants, such as a hot filament forming in a room-temperature matrix or an infinite planar simplified geometry. It was also studied in its transient evolution (for a review see e.g. [21]). In general, all of these thermal-based models were able to predict a threshold switching, but failed in the comparison with the experimental dependence of V_{th} on parameters such as the device length, with significant disagreement for sub-μm-thick films, or the temperature. The need to include an electronic component next to the thermal one, if not to substitute the second with the first, soon became clear, as more experimental evidences were found that a purely thermal mechanism could not fully explain switching-related phenomena (e.g. see the arguments based on the size of the on-state in [89] or on the polarity dependence of the switching time in [95], both derived from pulsed measurements). Furthermore, the predicted switching times do not seem compatible with those of modern highly scaled devices.
Nonetheless, the thermal aspect can have a cross influence on transport and switching properties. Bogoslovskiy and Tsendin formulated an electro-thermal model based on the Multi-phonon Tunneling Ionization (MTI) of negative-correlation-energy (U^-)

centers and on the lattice temperature increase due to the Joule dissipation during conduction [22, 20]. The existence of U^- centers is inferred due to considerations on the non-paramagnetic nature of chalcogenide glasses, while materials with a high concentration of occupied localized one-electron states should instead exhibit paramagnetism. To this regards, Anderson had proposed an effective attractive interaction between unpaired electrons that results in favoring electron pairing and diamagnetism of the material [11]. The U^- assumption can also account for other experimental facts such as the Fermi level pinning in chalcogenides and Stokes shift phenomena in photoluminescence spectroscopy [21]. These centers, here labeled D, can be found in the D^-, D^0 and D^+ states, having charge $-q$, 0 and $+q$, respectively. When a strong electric field F is applied, an electron e^- can be freed from a center by tunneling, in one of two ways, namely $D^- \rightarrow D^0 + e^-$ or $D^0 \rightarrow D^+ + e^-$. In the first case, the escaping carrier feels the action of a neutral trap potential (since self-interaction must not be considered). Analogously, in the second case it experiences a Coulombic interaction. These potential shapes can be used to calculate two different tunneling probabilities, showing the distinct potential-bending effects of an applied electric field F in the two situations. Lattice vibrations are included in the model through the Huang-Rhys approximation [5], and result in a temperature-dependent prefactor representing the thermal ionization probability. The reasoning above about electrons can be analogously done for holes, as it happens in chalcogenides. With all these ingredients, a conductivity $\sigma = \sigma(F)$ was calculated, assuming a constant mobility μ and incorporating the ionization effects in the hole concentration $p(T, F)$. At this point it was possible to write a thermal energy balance of the form

$$F^2 q p(T, F) \mu L = \lambda (T - T_0) \tag{3.7}$$

which is where the thermal component comes into play; L is the length of the device, λ a heat exchange constant and T_0 the ambient temperature. This equation was numerically solved, allowing one to find the switching behavior, with a fundamental role of the thermal part (see Figure 3.7).

A good qualitative agreement of the main experimental evidences is obtained with this model, but a quantitative one is lacking with respect to some observed trends, such as the dependence of the switching current density j_{th} on L ($j_{th} \propto L^{-0.7}$ instead of $j_{th} \propto L^{-1.4}$) and the T-dependence of F_{th}, which is too strong as compared with experiments. The authors ascribed these disagreements to possible spatial non-uniformities of T and F in the device, which the model does not account for.

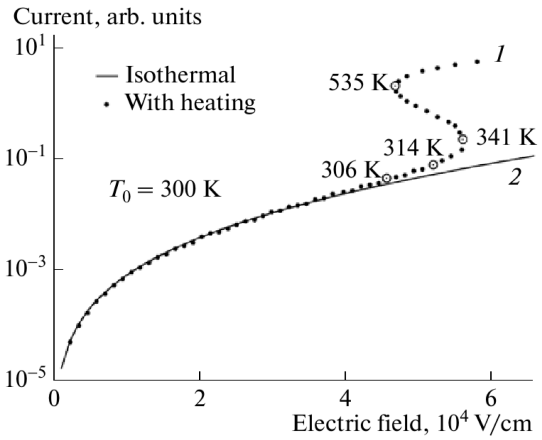

Figure 3.7: Calculated IV curve from the thermal-MTI model, with and without including the thermal component for the lattice. Figure taken from [21].

3.3 Summary

The rich variety of experimental evidences that characterizes PCM was described in this Chapter. This explains the huge number of theories devised along time to interpret as many of the observed features as possible. The nature of transport in chalcogenides seems to be an activated one, so that the temperature is one of the most important factors. Nonetheless, it is widely accepted that the switching phenomenon cannot be explained only on a thermal basis, and an electronic component must be considered. Whether some thermal effects are necessary or not is still a somewhat debated issue.

On the electronic side, several mechanisms have been proposed: trap-limited conduction following thermally activated emission, as in the Poole(-Frenkel) effect; generation-recombination dynamics, based on a field- and carrier-density-dependent phenomenon (such as impact ionization) to trigger a breakdown phenomenon; self-trapping effects resulting in high conductivity once a high charge density neutralizes the small-polaron conditions; or injection from the contacts together with the formation and neutralization of potential barriers.

Also, the possibility of a crystallization-based switching has been explored, in the assumption of an enhancement of the crystalline nucleation process by the electric field.

This variety of explanations and hypotheses, with sometimes equally valid comparison against experiments, highlight the difficulty of finding a really unifying theory, and ask for new efforts in trying to find one.

Part II

Study of electrical transport and switching properties

Chapter 4

Hydrodynamic-like modeling

Scaling is a central topic when it comes to phase-change memories [85, 59, 128, 131]. The shrinking of devices is both a phenomenal opportunity as well as a big issue to confront with [4, 136, 54, 34, 125]. It is of the utmost importance to acquire a sound understanding of the physics involved in PCM electronic behavior, and to use it to devise appropriate modeling tools.

As reviewed in Chapter 3, a number of approaches have been proposed to model the charge transport properties of amorphous chalcogenide switching devices, and shed light on the phenomenon of the Ovonic Threshold Switching. Some of the hypotheses made in the last decades include physical mechanisms such as Joule heating [86, 21], standard [93] or cooperative impact ionization [107], the formation of a conductive micrometric filament in the amorphous matrix due to generation-recombination dynamics [7, 8], the field-induced nucleation of a crystalline filament [116, 52], field-favored hopping among traps [25] eventually resulting in space-charge accumulation [24], and trap-limited conduction [47, 45]. For a comprehensive review, see e.g. [80].

All of these theories successfully model several different aspects of the chalcogenide transport behavior. However, they all rely on some level of approximation, and there is still room for improvement.

Thermally-induced switching actually predicts correct IV and switching behavior, but several quantitative dependences on the experimental parameters are not fully captured. The thermal switching times do not seem to be compatible with the OTS ones.

Impact ionization is questioned as an unlikely process, because of the typically low mean-free-path of carriers in disordered media, which keeps the energy gain from the field below the level needed for impact generation. Cooperative impact ionization overcomes this

limit, although the generation process requires an additional assumption on the mathematical form of the generation rate.
The formation of an electronic μm-size filament can be valid for big samples, but cannot be adapted to modern nano-devices [64].
The field-induced crystallization phenomenon reproduces several switching-related behaviors, but predictions about IV characteristic are not presented by the authors. Moreover, crystallization is not universally accepted as a requirement for the threshold switching.
Space-charge accumulation is evidenced by a Monte Carlo simulation, able to reproduce subthreshold conduction and switching. On the other side, somewhat ambiguous hypotheses on the electrostatics analysis and the choice of the transition mechanism (in this case hopping) were employed in the simulation.
The trap-limited approach treats the amorphous device as a quasi-1D object dominated by a Poole conduction mechanism: charge transport is studied at different energy levels, and a whole IV curve is predicted, in order to gain insight on the switching event. It overcomes several of the above-mentioned issues. However it finds an inconsistency between the electric potential and carrier distributions along the device (the latter assumed constant, thus neglecting diffusion of carriers).

In this Chapter, we describe a new quasi-1D model for charge transport in switching amorphous chalcogenides. The trap-limited picture of [47, 45] is taken as a starting point: carriers are trapped in localized states in the bandgap, from which they are excited above an activation energy (in a Poole-like fashion) so that they can freely transfer to another localization site and be retrapped. Direct hopping between traps is neglected since it is important only at low temperatures [76, 114], while thermally-assisted tunneling has a functional rate form similar to the Poole process and is thus unified with it. This allows to derive a first transport current-density vs field jF equation.
Carrier energy flux is also studied, including energy gains and losses, due to the interaction with the local electric field and to phonon emission mechanisms [49], and thus yielding a second equation for the energy balance at each slab of the device perpendicular to the transport direction z.
The model microscopically describes the device in out-of-equilibrium conditions with a local pseudo-Fermi level $E_F(z)$ and a local carrier temperature $T_e(z)$, together with the electric field $F(z)$. The introduction of the local carrier temperature stems from adding also the energy transport equation to the more common drift-diffusion framework.
The third equation needed to close the set of equations for such variables is provided by

the Poisson equation, which makes it possible to calculate the profiles of carrier concentration and energy (in terms of $E_F(z)$ and $T_e(z)$), together with the electric field along the device [2], analogously to the standard semiconductor Hydrodynamic approach [120]. Since the model is current-driven, these functions of z are found for a prescribed current-density, allowing one to calculate the device voltage drop as an output. The model is employed to study the subthreshold and switching behavior of switching chalcogenide devices [91]. Also, first-principle calculations can cast some light on the structural arrangement of amorphous chalcogenide [10, 29, 32], but the estimation of its consequences on localized electronic states is very challenging. Thus, the present model is employed also to explore several possible distributions of the trap energy levels in the bandgap [50] and estimate their impact on conduction.
The work is presented in the following through the inclusion of the two original papers [91, 50].

Hot-carrier trap-limited transport in switching chalcogenides

Enrico Piccinini,[1,a)] Andrea Cappelli,[2] Fabrizio Buscemi,[1] Rossella Brunetti,[2] Daniele Ielmini,[3] Massimo Rudan,[1] and Carlo Jacoboni[2]
[1]*"E. De Castro" Advanced Research Center on Electronic Systems (ARCES), Università di Bologna, Via Toffano 2/2, I-40125 Bologna, Italy*
[2]*Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Via Campi 213/A, I-41125 Modena, Italy*
[3]*Dipartimento di Elettronica e Informazione, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy*

(Received 1 August 2012; accepted 3 October 2012; published online 29 October 2012)

Chalcogenide materials have received great attention in the last decade owing to their application in new memory systems. Recently, phase-change memories have, in fact, reached the early stages of production. In spite of the industrial exploitation of such materials, the physical processes governing the switching mechanism are still debated. In this paper, we work out a complete and consistent model for transport in amorphous chalcogenide materials based on trap-limited conduction accompanied by carrier heating. A previous model is here extended to include position-dependent carrier concentration and field, consistently linked by the Poisson equation. The results of the new model reproduce the experimental electrical characteristics and their dependences on the device length and temperature. Furthermore, the model provides a sound physical interpretation of the switching phenomenon and is able to give an estimate of the threshold condition in terms of the material parameters, a piece of information of great technological interest. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4761997]

I. INTRODUCTION

With the recent introduction of the 22 nm node in the fabrication process, the semiconductor industry seems to be very close to its technological limit. According to the International Technology Roadmap for Semiconductors,[1] it could eventually be possible to scale down the actual devices to the next 1x generation only by redesigning either the device or the productive process, or both. It will be more and more complicated, if not impossible, to continue this trend much further.[1,2] Alternatively, new materials have to be explored and different architectural solutions implemented.

In the memory technology, non-silicon materials are being widely investigated in order to introduce faster, more scalable, and reliable devices. The present frontier of the research is represented by metal-oxide resistive RAMs and conductive-bridge RAMs,[1] while phase-change memory (PCM) prototypes have been studied in the last years[3] and are now in the early production stage.[4,5] Phase-change materials like chalcogenides are known from the early 1960s[6] and show the property of an easy, reversible transition between crystalline and amorphous phases, characterized by significant changes in optical reflectivity and electrical resistivity.[7] Due to high optical contrast between the two phases, these materials have been employed for optical storage since the mid 1990s.[8] In the new century, the strong difference in resistivity characterizing the two phases pushed chalcogenide materials as suitable candidates for solid-state nonvolatile memories. Furthermore, some chalcogenide glasses also feature an ovonic threshold-switching in the amorphous phase, which implies a negative differential resistance (NDR) in the current-voltage characteristic before the phase change takes place. Even though the first PCM arrays have already been released to the market, the knowledge of the physical process governing the ovonic switching mechanism is still a step behind. The availability of a theory able to identify the threshold point and predict the device behavior under given operating conditions is still sought by scientists and engineers to tailor the materials and the device set up.

The first microscopic interpretation of the switching behavior was due to Adler and coworkers,[9,10] who supposed the creation of a micrometer-wide low-resistance filament in the amorphous matrix, thus reducing the resistance of the device. Later on, the switching behavior was also found for sub-micrometer devices,[11] and this gave rise to alternative interpretations. Using the standard macroscopic quantities of the theory of transport in semiconductors like concentrations, velocities, and mobilities, it is possible to explain the switching in terms of drift-diffusion with impact ionization[12] or cooperative detrapping.[13]

Alternative interpretations are due to Karpov and coworkers,[14,15] and to Ielmini and Zhang,[16] and Ielmini.[17] The interpretation provided by Karpov and coworkers preserves the idea of conductive filaments as responsible for the switching behavior. From an energetic balance, they derive a model based on nucleation and growth of a thin crystalline filament that progressively expands in the amorphous matrix until it connects the electrodes, thus dramatically reducing the electrical resistance of the device.

On the other hand, Ielmini and coworkers proposed a thermally assisted trap-limited conduction mechanism, where

a)Electronic mail: enrico.piccinini@unimore.it.

the switching is ascribed to the increase in the average kinetic energy of the carriers through the device, as a result of the balance between the field-induced energy gain and the energy relaxation due to the scattering with phonons. A non-uniformity of the electric field is also found. This model has the capability of interpreting not only the electrical characteristic of a memory cell but also its dependence on thickness and temperature, which suggests a thermally activated conduction mechanism. Similarly, it has been shown through Monte Carlo simulations that a switching behavior can be triggered by a space-charge accumulation near the contacts, if a field-enhanced hopping conduction is considered.[18]

According to Ref. 17, a non-uniformity in the electric field across the device is found near and above the switching condition, which must be sustained by a non-negligible positive charge in the region close to the cathode. However, the equations used in Ref. 17 do not include the effect of the variable concentration of the active carriers across the device, which is instead approximated as a constant. In the present paper, we use the ideas presented in Ref. 17 as a starting point, and we work out a complete and consistent model which includes diffusion and achieves self-consistency between charge distribution and field. The results of this theoretical development reproduce the electrical characteristics and their dependences on the geometrical scaling factors and temperature, and provide a sound physical interpretation of the results.

II. THE MODEL

The presence of defects inherent to an amorphous material implies the existence of a number of localized trap states in the band gap. If the conduction was due to a pure hopping process, carriers would tunnel among traps until they reach the collecting contact. In the so called *trap-limited conduction* regime, instead, carriers undergo continuous trapping-detrapping processes by which they absorb and release energy in such a way that they overcome the potential barriers and move across the device. An intermediate mechanism (*thermally assisted tunneling*), where carriers absorb some energy and tunnel to the next trap, is also possible. The three mechanisms coexist, and the final transfer rate must encompass all of them. Since pure hopping is mainly effective in the low-temperature range (i.e., well below room temperature),[19,20] we can neglect it as we are not interested in that temperature range. As for thermally assisted tunneling, it was shown that its dependence upon temperature and field is similar to that of the trap-limited conduction process.[17] For these reasons, we develop the model for the trap-limited conduction case.

Let n_T be the trap concentration, assumed to be spatially uniform. Under equilibrium conditions, the concentration of electrons is independent of the position and given by the Fermi-Dirac statistics; the electron temperature T coincides with the lattice temperature T_0. In the off-equilibrium conditions, we assume translational symmetry along the x and y directions, so that the quantities of interest depend only on z, the current direction. The carrier concentration $n(z)$ is still described by a Fermi distribution where a quasi-Fermi level $E_F(z)$ replaces the Fermi level E_{F0} and $T(z)$ replaces T_0. A flat trap density of states $\Gamma = n_T/\Delta E_G$ is assumed inside the band gap $\Delta E_G = E_C(z) - E_V(z)$, where $E_C(z)$ and $E_V(z)$ denote the band gap edges. The band gap follows the potential profile along the z axis. The number of carriers between E_T and $E_T + \mathrm{d}E_T$ is given by

$$\mathrm{d}n = \frac{\Gamma}{1+\exp\left[\frac{E_T - E_F(z)}{kT(z)}\right]}\,\mathrm{d}E_T, \quad E_V(z) \le E_T \le E_C(z). \tag{1}$$

The integration of Eq. (1) over the band gap yields the carrier concentration at z

$$\begin{aligned} n(z) &= \int_{E_V(z)}^{E_C(z)} \frac{\Gamma}{1+\exp\left(\frac{E_T - E_F(z)}{kT(z)}\right)}\,\mathrm{d}E_T \\ &= n_T - \Gamma k T(z) \ln\left(\frac{1+\exp\frac{E_C(z)-E_F(z)}{kT(z)}}{1+\exp\frac{E_V(z)-E_F(z)}{kT(z)}}\right). \end{aligned} \tag{2}$$

The equilibrium value n_0 is obtained once $E_C(z)$, $E_V(z)$, $E_F(z)$, and $T(z)$ are replaced by their equilibrium values E_{C0}, E_{V0}, E_{F0}, and T_0

$$n_0 = n_T - \Gamma k T_0 \ln\left(\frac{1+\exp\frac{E_{C0}-E_{F0}}{kT_0}}{1+\exp\frac{E_{V0}-E_{F0}}{kT_0}}\right). \tag{3}$$

When E_{F0} is sufficiently far from the band edges E_{C0} and E_{V0} and close to midgap, $n_0 \approx n_T/2$. Under equilibrium conditions, the material is neutral. This situation may be obtained, for instance, by assuming donor-like (acceptor-like) traps and a negative (positive) compensating charge n_0.

The model is, in principle, the same if one deals with electrons or holes. For this reason, in the followings, we would rather term the two leads where carriers enter or leave the device as the *injecting* contact, $z = 0$, and the *collecting* contact, $z = \ell$, respectively. For the sake of simplicity, we develop and discuss the model only for the case of electrons.

Let Δz be the average traveled distance between the sites of successive detrapping-trapping events. Following Ref. 16, if the detrapping time is much longer than the traveling time, it can be taken equal to the transfer time τ. For thermally activated processes, τ is exponentially dependent on the barrier height experienced by the carriers at the detrapping event. Thus, we can define two different times $\tau_{\rightarrow}$ and $\tau_{\leftarrow}$ that apply to the motion in the two directions, i.e., from z to $z + \Delta z$ and from z to $z - \Delta z$. The local field gives rise to opposite effects, as shown below:

$$\tau_{\rightleftharpoons} = \tau_0 \exp\left[\frac{E_C(z) - E_T + \Delta U(z, z \pm \Delta z)}{kT_0}\right], \tag{4}$$

where τ_0 is a characteristic transfer time for the process at hand, and ΔU is the shift of the barrier height with respect to the equilibrium value due to the local electric field $F(z)$. With reference to Fig. 1, let $\lambda \Delta z$ be the distance of the maximum of the energy profile along the transition path from the trap having the lower z coordinate, with $0 < \lambda < 1$. Using a first-order approximation, the shifts ΔU result

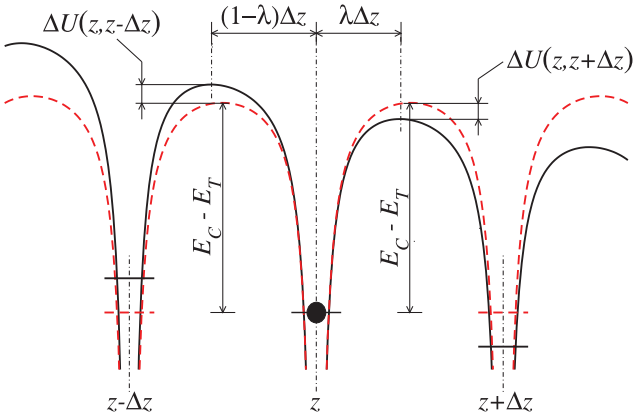

FIG. 1. Schematic view of the energy profile for transitions from z to $z+\Delta z$ and from z to $z-\Delta z$. The dashed profile corresponds to the equilibrium condition, while the solid curve is obtained under the influence of a local field $F(z)<0$ that enhances the transitions towards larger z's. Symbols are explained in the text.

$$\Delta U(z,z+\Delta z) = qF(z)\lambda\Delta z,$$
$$\Delta U(z,z-\Delta z) = -qF(z)(1-\lambda)\Delta z, \tag{5}$$

with q the absolute value of the electron charge. For the sake of simplicity, λ is assumed independent of z and equal to 1/2.

At a given position z, the carrier velocities in opposite directions around the energy E_T are given by

$$v_{\rightleftharpoons}(z,E_T) = \frac{\Delta z}{\tau_0}\exp\left[-\frac{E_C(z)-E_T}{kT_0}\right]\exp\left[\mp\frac{qF(z)\Delta z}{2kT_0}\right] \tag{6}$$

and, using Eq. (1), their average values over the entire distribution are

$$\langle v_{\rightleftharpoons}(z)\rangle = \frac{1}{n(z)}\int_{E_V(z)}^{E_C(z)}\frac{v_{\rightleftharpoons}(z,E_T)}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\Gamma\,\mathrm{d}E_T. \tag{7}$$

The current density $J(z)$ thus reads

$$J(z) = -q\left[n\left(z-\frac{\Delta z}{2}\right)\left\langle v_{\rightarrow}\left(z-\frac{\Delta z}{2}\right)\right\rangle - n\left(z+\frac{\Delta z}{2}\right)\left\langle v_{\leftarrow}\left(z+\frac{\Delta z}{2}\right)\right\rangle\right]. \tag{8}$$

By expanding the r.h.s of Eq. (8) to the first order in $\Delta z/2$, after some algebra we obtain

$$\frac{J(z)}{q} = -n(z)[\langle v_{\rightarrow}(z)\rangle - \langle v_{\leftarrow}(z)\rangle] + \frac{\Delta z}{2}\frac{\mathrm{d}}{\mathrm{d}z}\{n(z)[\langle v_{\rightarrow}(z)\rangle + \langle v_{\leftarrow}(z)\rangle]\}, \tag{9}$$

which, using Eqs. (1), (6), and (7), becomes

$$\frac{J(z)}{q} = -\frac{2\Delta z}{\tau_0}\sinh\left[-\frac{qF(z)\Delta z}{2kT_0}\right]A(z) + \frac{(\Delta z)^2}{\tau_0}\frac{\mathrm{d}}{\mathrm{d}z}\left\{\cosh\left[-\frac{qF(z)\Delta z}{2kT_0}\right]A(z)\right\}, \tag{10}$$

where

$$A(z) = \int_{E_V(z)}^{E_C(z)}\exp\left[-\frac{E_C(z)-E_T}{kT_0}\right]\frac{\Gamma}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\mathrm{d}E_T.$$

If J is fixed, as happens in the description of switching materials whose current-voltage characteristics are typically S-shaped, the model requires the determination of three unknown functions, namely, the electric field $F(z)$, the quasi-Fermi level $E_F(z)$, and the carrier temperature $T(z)$. Two additional equations to be coupled to Eq. (10) are then required.

One of them is the Poisson equation, which, due to Eqs. (2) and (3), reads

$$\frac{\mathrm{d}F(z)}{\mathrm{d}z} = \frac{\rho(z)}{\varepsilon} = -\frac{q}{\varepsilon}[n(z)-n_0] = \frac{q}{\varepsilon}\Gamma kT(z)\ln\left[\frac{1+\exp\frac{E_C(z)-E_F(z)}{kT(z)}}{1+\exp\frac{E_V(z)-E_F(z)}{kT(z)}}\left(\frac{1+\exp\frac{E_{V0}-E_{F0}}{kT_0}}{1+\exp\frac{E_{C0}-E_{F0}}{kT_0}}\right)^{\frac{T_0}{T(z)}}\right], \tag{11}$$

where $\varepsilon = \varepsilon_0\varepsilon_r$ is the dielectric constant of the material. In writing Eq. (11), we have taken into account that the material under equilibrium conditions must be neutral everywhere. Since E_{F0} has been assumed close to midgap and sufficiently far from the band edges, Eq. (11) simplifies into

$$\frac{\mathrm{d}F(z)}{\mathrm{d}z} \approx \frac{q}{\varepsilon}\left\{\Gamma kT(z)\ln\left[\frac{1+\exp\frac{E_C(z)-E_F(z)}{kT(z)}}{1+\exp\frac{E_V(z)-E_F(z)}{kT(z)}}\right] - \frac{n_T}{2}\right\}.$$

The third equation comes from the power balance. Depending on how effective electron-phonon scattering is in dissipating the power transferred to the carriers by the electric field, the average kinetic energy of the carriers (thus their temperature) may or may not stay tied to the equilibrium value. Let $\Delta E_{ex}^{TOT}(z)$ represent the *excess energy*, i.e., the difference between the actual energy of the carrier distribution in z and the energy that the same population would have if kept at the equilibrium temperature T_0

$$\Delta E_{ex}^{TOT}(z) = \int_{E_V(z)}^{E_C(z)}\frac{E_T-E_V(z)}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\Gamma\,\mathrm{d}E_T - \int_{E_V(z)}^{E_C(z)}\frac{E_T-E_V(z)}{1+\exp\left[\frac{E_T-\tilde{E}_F(z)}{kT_0}\right]}\Gamma\,\mathrm{d}E_T. \tag{12}$$

Here, $\tilde{E}_F(z)$ is defined at any z by imposing the constraint

$$\int_{E_V(z)}^{E_C(z)}\frac{1}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\mathrm{d}E_T = \int_{E_V(z)}^{E_C(z)}\frac{1}{1+\exp\left[\frac{E_T-\tilde{E}_F(z)}{kT_0}\right]}\mathrm{d}E_T,$$

which ensures the same population for the two distributions.

In order to write the power balance, one considers the power flowing through two different sections of the device at a distance $\mathrm{d}z$

$$\Phi(z+\mathrm{d}z) = \Phi(z) - J\mathrm{d}\varphi - \left.\frac{\partial \Delta E_{ex}^{TOT}(z)}{\partial t}\mathrm{d}z\right|_{\text{loss}}. \tag{13}$$

Here, $\Phi(z)$ is the energy density flux in z; $\mathrm{d}\varphi = -F(z)\mathrm{d}z$ is the variation of the electrostatic potential in the z direction, and the last term of the r.h.s. represents the power exchanged via inelastic electron-phonon scattering. The derivative can be expressed in the relaxation-time approximation as[21]

$$\left.\frac{\partial \Delta E_{ex}^{TOT}(z)}{\partial t}\right|_{\text{loss}} = \frac{\Delta E_{ex}^{TOT}(z)}{\tau_r},$$

τ_r being a relaxation time constant. After expanding the l.h.s. of Eq. (13) to the first order in $\mathrm{d}z$ and dividing both sides by $\mathrm{d}z$, one gets

$$\frac{\mathrm{d}\Phi(z)}{\mathrm{d}z} = JF(z) - \frac{\Delta E_{ex}^{TOT}(z)}{\tau_r}. \tag{14}$$

The energy density flux $\Phi(z)$ can be calculated following the same scheme adopted for the current density $J(z)$ in Eq. (8)

$$\Phi(z) = n\left(z-\frac{\Delta z}{2}\right)\left\langle P_{\rightarrow}\left(z-\frac{\Delta z}{2}\right)\right\rangle - n\left(z+\frac{\Delta z}{2}\right)\left\langle P_{\leftarrow}\left(z+\frac{\Delta z}{2}\right)\right\rangle, \tag{15}$$

where

$$\langle P_{\rightleftharpoons}(z)\rangle = \frac{1}{n(z)}\int_{E_V(z)}^{E_C(z)} v_{\rightleftharpoons}(z,E_T)\frac{E_T - E_V(z)}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\Gamma\,\mathrm{d}E_T \tag{16}$$

represent the two average energy fluxes of the carrier distribution in opposite directions at a given coordinate z.

The r.h.s. of Eq. (15) can be replaced with its first-order approximation in $\Delta z/2$, this leading to

$$\Phi(z) = n(z)[\langle P_{\rightarrow}(z)\rangle - \langle P_{\leftarrow}(z)\rangle] - \frac{\Delta z}{2}\frac{\mathrm{d}}{\mathrm{d}z}\{n(z)[\langle P_{\rightarrow}(z)\rangle + \langle P_{\leftarrow}(z)\rangle]\}.$$

Equation (14) now reads

$$\frac{\mathrm{d}}{\mathrm{d}z}\{n(z)[\langle P_{\rightarrow}(z)\rangle - \langle P_{\leftarrow}(z)\rangle]\} - \frac{\Delta z}{2}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\{n(z)[\langle P_{\rightarrow}(z)\rangle + \langle P_{\leftarrow}(z)\rangle]\} = JF(z) - \frac{\Delta E_{ex}^{TOT}(z)}{\tau_r}, \tag{17}$$

or, using Eqs. (1), (6), (12), and (16),

$$\frac{2\Delta z}{\tau_0}\frac{\mathrm{d}}{\mathrm{d}z}\left\{\sinh\left[-\frac{qF(z)\Delta z}{2kT_0}\right]B(z)\right\} - \frac{(\Delta z)^2}{\tau_0}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left\{\cosh\left[-\frac{qF(z)\Delta z}{2kT_0}\right]B(z)\right\} = JF(z) - \frac{1}{\tau_r}\left\{\int_{E_V(z)}^{E_C(z)}\Gamma\frac{E_T-E_V(z)}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\mathrm{d}E_T - \int_{E_V(z)}^{E_C(z)}\Gamma\frac{E_T-E_V(z)}{1+\exp\left[\frac{E_T-\tilde{E}_F(z)}{kT_0}\right]}\mathrm{d}E_T\right\}, \tag{18}$$

where

$$B(z) = \int_{E_V(z)}^{E_C(z)}\Gamma\exp\left[-\frac{E_C(z)-E_T}{kT_0}\right]\frac{E_T-E_V(z)}{1+\exp\left[\frac{E_T-E_F(z)}{kT(z)}\right]}\mathrm{d}E_T.$$

The set of Eqs. (10), (11), and (18) leads to the determination of the unknown functions $F(z)$, $E_F(z)$, and $T(z)$ for any given current density J. Equation (18) involves the second derivatives of the unknown functions; nevertheless, a numerical analysis has shown that the term proportional to the second derivative in Eq. (18) is negligible with respect to the other term, but for a narrow region close to the injecting contact at the highest currents. For the latter case, this term, though effective to some extent, still remains smaller than the other one in the l.h.s. of Eq. (18). The second-derivative contribution can thus be neglected without substantially affecting the physical results.

However, an analytical closed form for the solution of the set of Eqs. (10), (11), and (18) cannot be obtained because the integrals in Eq. (18) have to be evaluated numerically. One can overcome this problem by replacing the Fermi distribution function with a suitable approximation $\chi(E_T,z)$ defined in such a way that: (i) $\chi(E_F(z)) = 1/2$, (ii) $\chi(E_T,z)$ shares the same asymptotical values of the original Fermi-Dirac distribution, and (iii) the following symmetry holds $1-\chi(E_F-\Delta) = \chi(E_F+\Delta)$. The above requirements are satisfied for instance by

$$\chi(E_T,z) = \begin{cases} 1-\frac{1}{2}\exp\left[\Omega\frac{E_T-E_F(z)}{kT(z)}\right] & \text{if } E_T < E_F(z) \\ \frac{1}{2}\exp\left[-\Omega\frac{E_T-E_F(z)}{kT(z)}\right] & \text{if } E_T \geq E_F(z) \end{cases}, \tag{19}$$

where a good choice for the parameter Ω is $\Omega = 3/4$, as shown in Fig. 2.

Let us introduce the following dimensionless functions:

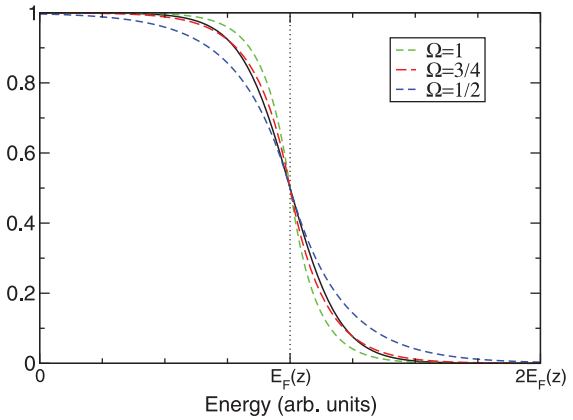

FIG. 2. Comparison between the Fermi distribution function (solid dotted line) and the approximating function $\chi(E_T, z)$ for different values of the parameter Ω (dashed lines). The best approximation is obtained when $\Omega = 3/4$.

$$f(z) = -\frac{qF(z)\Delta z}{2kT_0}, \quad g(z) = \frac{E_F(z) - E_{F0}}{kT_0}, \quad t(z) = \frac{T(z)}{T_0},$$

which describe the off-equilibrium local field, the shift of the quasi-Fermi level, and the electron temperature, respectively.

After introducing the $\chi(E_T, z)$ function in place of the Fermi distribution, and neglecting the second derivative in Eq. (18), the set given by Eqs. (10), (11), and (18) can be manipulated through the straightforward, though lengthy, calculations summarized in the Appendix to yield

$$\frac{df}{dz} = N^*(g,t), \tag{20}$$

$$J_g(f,g,t)\frac{dg}{dz} + J_t(f,g,t)\frac{dt}{dz} = J^*\left(f,g,t,\frac{df}{dz}\right), \tag{21}$$

$$H_g(f,g,t)\frac{dg}{dz} + H_t(f,g,t)\frac{dt}{dz} = H^*\left(f,g,t,\frac{df}{dz}\right). \tag{22}$$

From the above, one immediately obtains

$$\frac{df}{dz} = N^*, \quad \frac{dg}{dz} = \frac{J^*H_t - J_tH^*}{J_gH_t - J_tH_g}, \quad \frac{dt}{dz} = \frac{J_gH^* - J^*H_g}{J_gH_t - J_tH_g}. \tag{23}$$

The definitions of the symbols can be found in the Appendix.

Using a first-order Runge-Kutta integration scheme, the equations above are numerically solved once the values of $f(z)$, $g(z)$, and $t(z)$ are provided at the coordinate $z = 0$. Two boundary conditions can easily be inferred by supposing that the electrons at the injecting contact are at equilibrium, namely, $E_F(0) = E_{F0}$ and $T(0) = T_0$. The definitions of $g(z)$ and $t(z)$ allow for the direct conversion of these boundary conditions into $g(0) = 0$ and $t(0) = 1$.

The value for $f(z)$ at the boundary must be such that the global charge neutrality of the system holds true

$$\varepsilon[F(0) - F(\ell)] + q\int_0^\ell [n(z) - n_0]\,dz = 0. \tag{24}$$

A trial-and-error procedure for the initial value $f(0)$ is applied, until the solutions for $f(z)$, $g(z)$, and $t(z)$ satisfy Eq. (24) within the desired precision.

III. RESULTS AND DISCUSSION

A. Current-voltage characteristics

We report in Fig. 3 the $I(V)$ characteristic obtained with the model given by Eqs. (20)–(22) (solid black curve), where V is calculated as the integral of the field $F(z)$ along the device. For comparison, the result in Ref. 17 (dash-dotted red curve) and experimental data for a GST-225 memory cell available therein are also shown (dots). These data correspond to a memory cell with a bottom contact electrode with cross-section $\Sigma = 1000\,\text{nm}^2$ and length $\ell = 40\,\text{nm}$. The dashed green curve refers to a calculation performed using the present model with the same parameters reported in Ref. 17. It is clearly seen that, for the same set of parameters, the new model would lead to an evident increase in the conductivity (the dashed curve yields a given current at a lower voltage) and to a reduction of the threshold voltage, without substantially affecting the threshold current. These differences are due to the different approximation adopted for the carrier distribution function by the two models. While in Ref. 17 the carriers that contribute to transport are only those above the Fermi level, in the present model the Fermi distribution function is integrated over the entire band gap, thus making the conductivity higher. Since the threshold point is determined mainly by the current, as will be discussed later on, the increased conductivity implies a smaller threshold voltage.

In order to fit the experimental data, it has then been necessary to calibrate the parameters. The new parameter set is reported in Table I along with the set used in Ref. 17 for the sake of comparison. The band gap considered here is compatible with literature data for the amorphous GST-225. The two time constants τ_0 and τ_r are consistent with those suggested by Mott and Davis for amorphous semiconductors.[19]

The agreement between the results of the present model with the new set of parameters and the experimental data is quite good. Differences with the results of model of Ref. 17

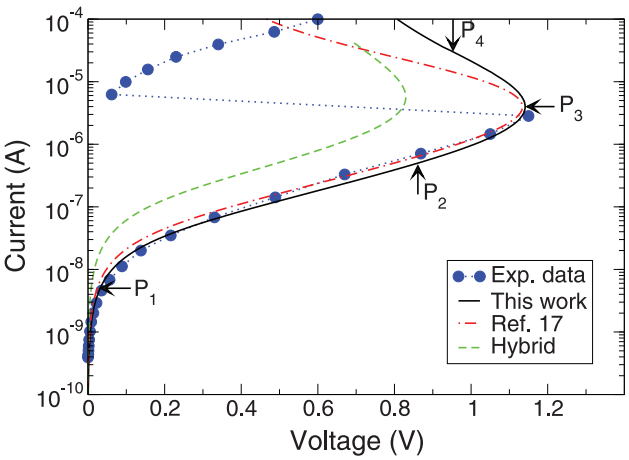

FIG. 3. $I(V)$ characteristic obtained for the present model compared to experimental data and to the corresponding curve taken from Ref. 17. The dashed green curve refers to a "hybrid" configuration where the parameters proposed in Ref. 17 have been used in the present model. The arrows show the positions along the $I(V)$ curve of the points $P_1 \ldots P_4$ cited in the text and in Fig. 4.

TABLE I. Parameters used for the best fit reported in Fig. 3. The values used in Ref. 17 are also reported for comparison.

Parameter	Present work	Ref. 17
ΔE_G	0.68 eV	0.6 eV
$n_T/\Delta E_G$	10^{20} cm^{-3} eV^{-1}	10^{20} cm^{-3}eV^{-1}
τ_0	1.2×10^{-14} s	1.0×10^{-14} s
τ_r	0.78×10^{-13} s	1.0×10^{-13} s
Δz	7×10^{-7} cm	7×10^{-7} cm
ε_r	15	15
T_0	298 K	298 K

are found in the NDR region, where the present model estimates larger potential drops for any given current. We point out, however, that above the switching point, the experimental data depend on the characteristics of the external circuit. For this reason, they should not be considered as for the sake of comparison. In fact, under different experimental conditions the current may rise nearly vertically at a holding voltage.[22] Then, the crystallization of the material occurs. Results from the present model obtained far above threshold should thus be considered only qualitatively.

B. Microscopic interpretation of the switching

In order to analyze the microscopic process leading to the threshold switching, it is useful to consider the physical quantities of interest along the device. Apart from the amplitude of the variations, a common behavior for the dimensionless electric field $f(z)$ and carrier temperature $t(z)$, and for the carrier concentration $n(z)$, can be outlined. The calculated profiles for these quantities are reported in Fig. 4 for the four points $P_1 \ldots P_4$ shown in Fig. 3.[23]

Since the injecting contact acts as an infinite reservoir of carriers in thermal equilibrium with the lattice, carriers enter the device at thermal equilibrium. However, due to the imbalance between the power provided by the field and the power loss due to electron-phonon scattering, they tend to heat up as long as they travel along the device. This effect is negligibly small at the lowest currents. The increase in the carrier temperature would enhance the flux towards the collecting contact and, in order to keep J at its prescribed value, it is compensated by a decrease in the carrier concentration and in the electric field, consistently with the Poisson equation. Thus, a positive charge accumulates near the injecting contact. The position of the quasi-Fermi level shifts towards a lower energy to account for such a change in the carrier concentration.

Going farther from the injecting contact, the heated carrier population enables a more effective dissipation through inelastic scattering, which prevails over the power provided by the electric field. The electron temperature, after reaching a maximum value, decreases. A thermal overshoot is thus created near the injecting contact. Since the charge must vary continuously in space, the presence of a positive charge implies that the electric field continues to decrease also in the region where the carrier temperature is reduced, although at a lower rate. In order to preserve the current, the carrier

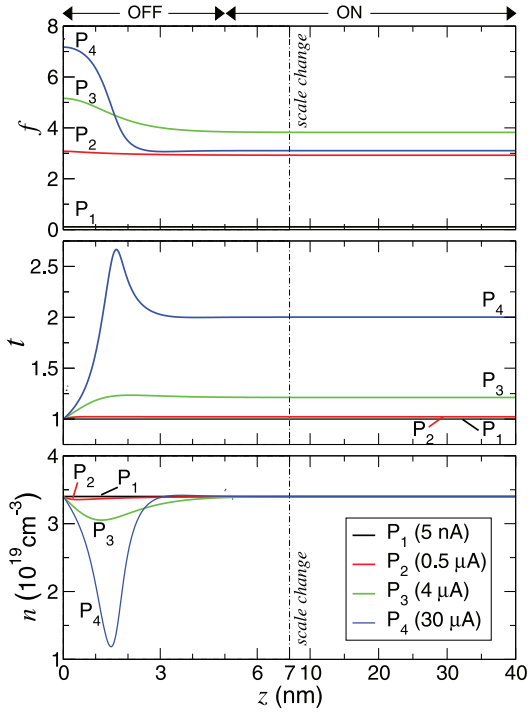

FIG. 4. Dimensionless electric field $f(z)$ (top), dimensionless electron temperature $t(z)$ (middle), and electron concentration $n(z)$ (bottom) for the four currents identified in Fig. 3. Note the scale change between the OFF and the ON regions at $z = 7$ nm.

concentration must compensate the reduction in velocity due to the smaller field and temperature, so that it increases significantly and approaches again the equilibrium value. When the charge neutrality is attained far enough from the injecting contact, the electric field and the carrier temperature saturate as well. It follows that a saturation value for the position of the quasi-Fermi level is also found.

As in Ref. 16, we refer to the OFF region as the zone of length ℓ_{OFF} close to the injecting contact where the physical quantities vary most appreciably, and to the ON region as the longer zone of length ℓ_{ON} where they have their saturation values, as shown in Fig. 4.

Even though the interpretation given above applies at any current, in the first part of the characteristic (point P_1, $I = 5$ nA) the electric field is low and does not provide enough power to determine an appreciable electron heating. This fact implies that the electron temperature and concentration are tied to their equilibrium values across the whole device (Ohmic behavior).

A similar situation applies also for most part of the subsequent exponential region. However, as the current increases, the electric field $f(0)$ grows rapidly (P_2, $I = 0.5\,\mu$A) and eventually triggers an appreciable electron heating. In turn, the latter causes a slight depletion of the OFF region and a corresponding weak non-uniformity of the electric field. Due to the exponential relationship linking the current and the field, this picture is more and more evident as the switching current is approached (P_3, $I = 4\,\mu$A).

As the current is increased above the threshold point (P_4, $I = 30\,\mu$A), $f(0)$ still continues to grow, thus inducing a

larger carrier heating in the OFF region. As a consequence, the local electric field rapidly falls, and the electron concentration is strongly depleted to keep the current constant. The high value of the electron temperature in the OFF region reflects into a high saturation value also in the ON region, which is the key condition to restore the equilibrium between energy gain and energy relaxation. Since in the ON region the carrier concentration has always the equilibrium value at any current, every further increase in the carrier temperature can only be compensated by a further decrease in the electric field. For such a reason, the resulting electric field is smaller than that found at the switching point. Since the ON region is substantially much longer than the OFF region, a smaller field in the ON region leads also to smaller potential drop across the device.

Finally, in the region above the switching point, the power density dissipated and transferred to the lattice via electron-phonon scattering is high. This may give rise to lattice heating, which is a favorable condition for the creation of a local crystalline nucleus that can eventually evolve into a crystalline filament. The incorporation of the Fourier heat equation and a local lattice temperature in the model goes beyond the scope of the present paper and is planned for a future work.

In conclusion, few words must be spent on the role of the Poisson equation, by comparing the present results with those of Ref. 17. As in the ON region charge neutrality is kept, both models provide similar results; on the contrary, in the OFF region the introduction of a self-consistent non-uniform carrier distribution yields a more accurate physical picture. In particular, the presence of a minimum in the carrier concentration profile implies a thermal overshoot and an initial different curvature of the field profile.

C. The switching condition

The last part of this section is devoted to the analysis of the switching condition as a function of the device length (Fig. 5) and the lattice temperature (Fig. 6).

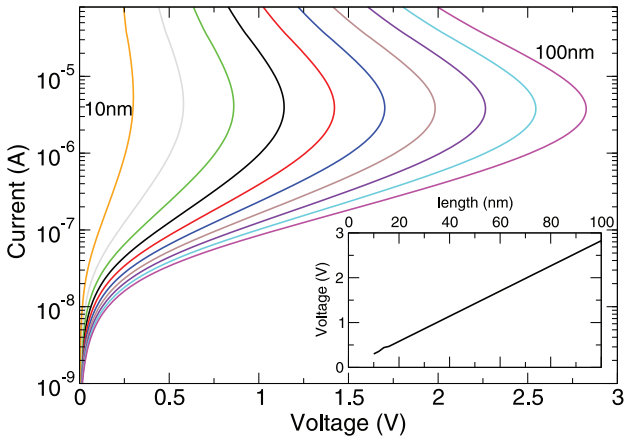

FIG. 5. $I(V)$ characteristics for devices with different lengths ranging from 10 nm to 100 nm with steps of 10 nm. The inset shows the linear relationship between the threshold voltage and the device length.

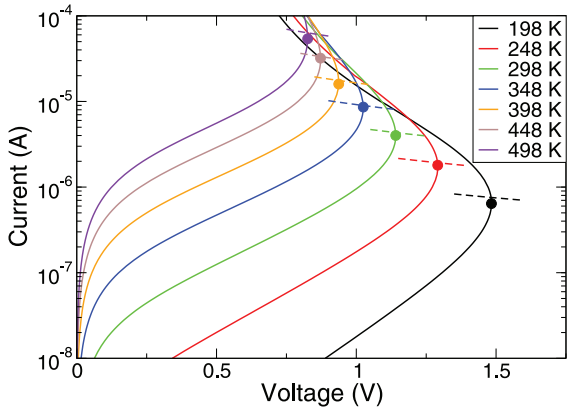

FIG. 6. $I(V)$ characteristics for different lattice temperatures in the range 198 K to 498 K. The intersections with the dashed lines represent the switching points as calculated from Eq. (27).

The OFF region, as shown in Fig. 4, extends over about the first 5 nm from the injecting contact, which is the space where carriers cannot fully relax the power provided by the field through electron-phonon scattering and heat up. The microscopic phenomena occurring in the OFF region suggest that ℓ_{OFF} must primarily depend on material properties, like the density of traps n_T, the position of the Fermi level with respect to the bottom of the conduction band, and the relaxation time τ_r, but must be independent of the device length. As a consequence, when the latter exceeds approximately $2\ell_{OFF}$, the potential drop in the ON region dominates over that in the OFF region, and the threshold voltage scales almost linearly with the device length, as shown in the inset of Fig. 5. In this case, the error made by considering the electric field F_{th} in the ON region as representative of the field in the entire device at threshold is negligible. On the other hand, small deviations from linearity are found for shorter devices, as the potential drop in the OFF region gains relative importance over that in the ON region. According to the present model, ultra-short devices are not expected to show a NDR portion of the $I(V)$ characteristic, but, rather, a steep rise of the current with an almost constant potential, in agreement with the results of Fig. 5.

In order to assess the effect of the lattice temperature on the switching condition, we report in Fig. 6 the $I(V)$ characteristics for different lattice temperatures. It is found that the threshold current increases with temperature, while the threshold potential decreases, as typical of chalcogenide glasses.[24] Moreover, the calculations show that the threshold potentials tend to accumulate towards a limiting value as the temperature increases, suggesting the existence of a minimum threshold electric field.

With the purpose of understanding this phenomenon, we recall first that the ON region is defined as the zone where the equilibrium carrier concentration is restored, and the electric field and carrier temperature saturate. This physical condition is expressed by making the l.h.s. of Eq. (17) to vanish; viz.,

$$JF(z) - \frac{\Delta E_{ex}^{TOT}(z)}{\tau_r} = 0.$$

Taking F_{th} as representative of the field within the device, the product $J_{th}F_{th}$ is the input power density w_{th} at the threshold point. Here and in the following, the suffix *th* indicates the quantities evaluated at the threshold point. The definitions of $\Delta E_{ex}^{TOT}(z)$ and $U(t,\alpha,\beta)$ given by Eqs. (A4) and (A9) in the Appendix yield

$$w_{th} = \frac{I_{th}}{\Sigma}\frac{\Delta\varphi_{th}}{\ell} = \Gamma\,(kT_0)^2\,\frac{t_{th}^2 U(t,\alpha,\beta)|_{th} - U_0(\alpha_0',\beta_0')|_{th}}{\tau_r}.$$

If $t_{th} < 2.5$, as happens for a wide range of lattice temperatures up to 600 K (see also Fig. 7), the exponents in $U(t,\alpha,\beta)$ can be neglected, and the above equation further simplifies to

$$w_{th} = \frac{16}{9}\,\Gamma k^2\,\frac{T_{th}^2 - T_0^2}{\tau_r}, \tag{25}$$

where the condition that the quasi-Fermi level in the ON region is close to the local equilibrium value ($g_{th} \approx 0$) has also been used. The lack of an analytical solution for $t(z)$ makes it impossible to further simplify Eq. (25) into a compact form showing the dependences on the parameters of the material only. However, by means of numerical analyses, it is possible to obtain an empirical formula linking t_{th} with the lattice temperature.

Since the threshold current strongly increases with temperature overcompensating the decrease in the threshold potential, the input power at threshold increases indefinitely. The above considerations suggest an approximate function for t_{th} like, e.g.,

$$t_{th} = \frac{\theta T_0 - T^*}{T_0 - T^*}, \tag{26}$$

where the parameter T^* has the physical meaning of an asymptotical temperature up to which the NDR regions of the $I(V)$ characteristics are possible. For the case at hand, by best fitting the data reported in Fig. 7 (dashed green line), one finds $\theta \approx 2/3$ and $T^* \approx 730$ K, a value slightly larger than the glass transition temperature of the material. We point out, however, that these values should be considered with care, as they depend non-linearly on the parameters of the model. Numerical calculations have shown that a major role is played by the energy relaxation time τ_r, which significantly influences the threshold current and voltage. In fact, increasing the energy relaxation time by one order-of-magnitude reduces the asymptotical temperature by about 100 K, whereas reducing the relaxation time by one order-of-magnitude let the asymptotical temperature raise only by 35 K. The θ coefficient ranges instead from 0.75 down to 0.63 for the same variations of τ_r.

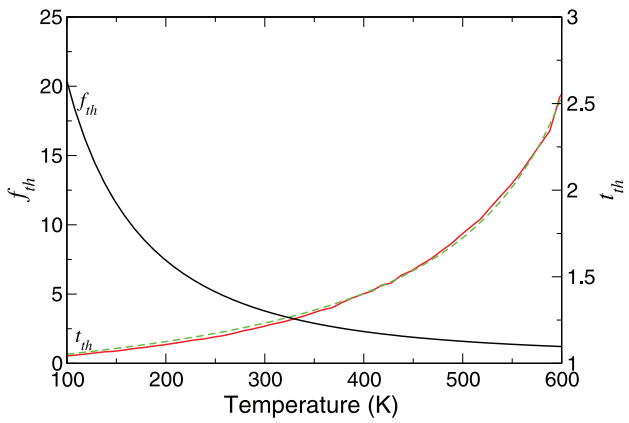

FIG. 7. Values at the threshold point of the dimensionless electric field and of the carrier temperature. The dashed green line is the best-fit approximation of t_{th} obtained by means of Eq. (26).

By inserting Eq. (26) into Eq. (25), one finds

$$w_{th} = \frac{16\,n_T\,(kT_0)^3}{9\,\tau_r\,\Delta E_G}\,\frac{(\theta+1)kT_0 - 2kT^*}{(kT_0 - kT^*)^2}(\theta - 1) \tag{27}$$

that represents the critical power at threshold as a function of the lattice temperature. In order to provide an estimate of the switching point, this formula must be combined with Eq. (A2) that expresses the current in the ON region. Taking again into account that the quasi-Fermi level in the ON region is close to its local equilibrium value, and using Eq. (26) to express the carrier temperature at threshold, it is possible to obtain the threshold current as a function uniquely of the dimensionless field f_{th}. By inserting the expression for J_{th} from Eq. (A2) into Eq. (27), after some algebra, one gets

$$f_{th}\sinh(f_{th}) = \frac{8}{9}\frac{\tau_0}{\tau_r}\exp\left(\frac{\Delta E_G}{2kT_0}\right)kT_0\,\frac{(\theta+1)kT_0 - 2kT^*}{(kT_0 - kT^*)^2} \times \frac{\theta - 1}{Q_{th}(T_0, T^*, \theta)}, \tag{28}$$

where $Q_{th}(T_0, T^*, \theta)$ is calculated by setting $t = t_{th}$ and $g = 0$ in Eq. (A8). Once f_{th} is known from Eq. (28), the threshold potential and current are given as $\Delta\varphi_{th} = F_{th}\ell$ and $J_{th} = w_{th}/F_{th}$, respectively. The above results allow for a rough estimate of the threshold point from the physical parameters of a given switching material.

IV. CONCLUSIONS

An enhanced model for transport in amorphous chalcogenides has been worked out to achieve self-consistency between the electric field present along the device and the local density and energy distribution of carriers. This characteristic was, in fact, missing in the models available in the literature.

The model relies on the hypothesis that conduction can be described by means of trap-limited transport, i.e., sequences of detrapping events letting a carrier move above the conduction-band mobility gap, followed by fast recaptures by different traps. Within this framework, it is possible to write two equations for the charge and the energy fluxes where two characteristic times are present: the detrapping time and the energy relaxation time. A third equation expresses charge-field consistency. The carrier concentration and the charge and energy fluxes are obtained by means of the integration of the Fermi distribution function over the

band gap, which is the energy region filled by trap states. The solution of the above set of equations yields the three main physical quantities that represent the unknowns of the problem, namely, the carrier concentration, the carrier temperature, and the electric field as functions of the position along the device.

After a proper identification of the physical parameters, the calculated current-voltage characteristics are in good agreement with experimental data. Moreover, the model correctly identifies both the temperature and device-length dependences of the $I(V)$ curves.

The ovonic switching behavior of chalcogenide glasses is explained by the effect of carrier heating. In trap-limited conduction, conductivity is determined mainly by the detrapping time, which depends on the energy barrier separating the energy of the carrier sitting in a trap from the conduction-band mobility edge. Detrapping processes are therefore favored by a lowering of the barrier due to the electric field and by the raising of the carrier energy due to their heating. At and above threshold, any further increase in the current requires a higher carrier temperature, which reduces the field in the largest part of the device, so that the potential drop decreases. The model can thus predict the switching point starting from the physical parameters of the material, which is a key achievement for technological applications.

ACKNOWLEDGMENTS

Part of this work has been carried out under the Contract No. 3477131/2011 of the Intel Corporation whose support is gratefully acknowledged.

APPENDIX: ANALYTICAL DERIVATION OF THE CONSTITUTIVE EQUATIONS

This appendix contains in some details the calculations which transform Eqs. (10), (11), and (18) in the main text into Eqs. (20)–(22) so that the interested reader can follow the mathematical development between the two sets of equations.

Let us introduce the dimensionless auxiliary variables

$$\alpha_0(z) = C_0 + g(z), \quad \alpha(z) = \frac{\alpha_0}{t(z)}, \quad \beta_0(z) = C_0 - g(z),$$

$$\text{and} \quad \beta(z) = \frac{\beta_0}{t(z)},$$

with $C_0 = \Delta E_G/2kT_0$. Following the calculations sketched in Sec. II, after replacing the Fermi-Dirac distribution with $\chi(E_T, z)$, one finds (omitting the indication of the dependence on z of the unknown functions)

$$n = \int_{E_V}^{E_C} \Gamma\chi(E_T, z)\, \mathrm{d}E_T = \Gamma k T_0\, tN(\alpha, \beta), \tag{A1}$$

$$n(\langle v_{\rightarrow}\rangle - \langle v_{\leftarrow}\rangle) = \int_{E_V}^{E_C} \Gamma\chi(E_T, z)\frac{\Delta z}{\tau_0}\exp\left(-\frac{E_C - E_T}{kT_0}\right)\sinh\left(-\frac{qF\Delta z}{2kT_0}\right)\mathrm{d}E_T$$
$$= \frac{\Gamma k T_0 \Delta z}{\tau_0}\sinh(f)\exp(-\beta_0)Q(t, \alpha_0, \alpha, \beta_0, \beta), \tag{A2}$$

$$n(\langle v_{\rightarrow}\rangle + \langle v_{\leftarrow}\rangle) = \int_{E_V}^{E_C} \Gamma\chi(E_T, z)\frac{\Delta z}{\tau_0}\exp\left(-\frac{E_C - E_T}{kT_0}\right)\cosh\left(-\frac{qF\Delta z}{2kT_0}\right)\mathrm{d}E_T$$
$$= \frac{\Gamma k T_0 \Delta z}{\tau_0}\cosh(f)\exp(-\beta_0)Q(t, \alpha_0, \alpha, \beta_0, \beta), \tag{A3}$$

$$\Delta E_{ex}^{TOT} = \int_{E_V}^{E_C} \Gamma[\chi(E_T, z) - \tilde{\chi}(E_T, z)|_{T=T_0}](E_T - E_V)\,\mathrm{d}E_T = \Gamma(kT_0)^2\,[t^2 U(t, \alpha, \beta) - U_0(\alpha_0', \beta_0')], \tag{A4}$$

$$n(\langle P_{\rightarrow}\rangle - \langle P_{\leftarrow}\rangle) = \int_{E_V}^{E_C} \Gamma\chi(E_T, z)(E_T - E_V)\frac{\Delta z}{\tau_0}\exp\left(-\frac{E_C - E_T}{kT_0}\right)\sinh\left(-\frac{qF\Delta z}{2kT_0}\right)\mathrm{d}E_T$$
$$= \frac{2\Gamma\Delta z(kT_0)^2}{\tau_0}\sinh(f)\exp(-\beta_0)S(t, \alpha_0, \alpha, \beta_0, \beta), \tag{A5}$$

$$n(\langle P_{\rightarrow}\rangle + \langle P_{\leftarrow}\rangle) = \int_{E_V}^{E_C} \Gamma\chi(E_T, z)(E_T - E_V)\frac{\Delta z}{\tau_0}\exp\left(-\frac{E_C - E_T}{kT_0}\right)\cosh\left(-\frac{qF\Delta z}{2kT_0}\right)\mathrm{d}E_T$$
$$= \frac{2\Gamma\Delta z(kT_0)^2}{\tau_0}\cosh(f)\exp(-\beta_0)S(t, \alpha_0, \alpha, \beta_0, \beta) \tag{A6}$$

with

$$N(\alpha,\beta)=\alpha+\frac{2}{3}\exp\left(-\frac{3}{4}\alpha\right)-\frac{2}{3}\exp\left(-\frac{3}{4}\beta\right), \tag{A7}$$

$$\begin{aligned}Q(t,\alpha_0,\alpha,\beta_0,\beta)&=\frac{18}{9-16t^2}-\exp(-\alpha_0)\left[2-\frac{4t}{4t+3}\exp\left(-\frac{3}{4}\alpha\right)\right]\\&+\frac{4t}{4t-3}\exp\left(\beta_0-\frac{3}{4}\beta\right),\end{aligned} \tag{A8}$$

$$\begin{aligned}U(t,\alpha,\beta)&=\frac{16}{9}+\frac{\alpha^2}{2}-\frac{8}{9}\exp\left(-\frac{3}{4}\alpha\right)\\&-\frac{4}{3}\left(\frac{C_0}{t}+\frac{2}{3}\right)\exp\left(-\frac{3}{4}\beta\right),\end{aligned} \tag{A9}$$

and

$$\begin{aligned}S(t,\alpha_0,\alpha,\beta_0,\beta)&=\alpha_0\frac{9}{9-16t^2}+9\frac{48t^2-9}{(9-16t^2)^2}\\&+\exp(-\alpha_0)\left[1-\frac{1}{2}\left(\frac{4t}{4t+3}\right)^2\exp\left(-\frac{3}{4}\alpha\right)\right]\\&+\frac{4t}{4t-3}\left(C_0-\frac{2t}{4t-3}\right)\exp\left(\beta_0-\frac{3}{4}\beta\right).\end{aligned} \tag{A10}$$

The $U_0(\alpha_0',\beta_0')$ function in Eq. (A4) can be calculated by setting $(t,\alpha,\beta)=(1,\alpha_0',\beta_0')$ in $U(t,\alpha,\beta)$, with α_0' and β_0' fulfilling the constraints $tN(\alpha,\beta)=N(\alpha_0',\beta_0')$ and $\alpha_0'+\beta_0'=2C_0$.

Under equilibrium, it is $f=0$, $g=0$, and $t=1$. As a consequence, $\alpha_{eq}=\beta_{eq}=C_0$. The equilibrium electron concentration thus reads

$$n_0=\Gamma kT_0C_0=\frac{n_T}{2}.$$

By means of this relationship and of Eq. (A1), Eq. (11) can then be recast in terms of the dimensionless auxiliary variables as follows:

$$\frac{\mathrm{d}f}{\mathrm{d}z}=C_1[tN(\alpha,\beta)-C_0] \tag{A11}$$

with $C_1=q^2\Gamma\Delta z/(4\varepsilon)$.

Equation (10) can be approximated by means of Eqs. (9), (A2), and (A3). Recalling that

$$\frac{\mathrm{d}\alpha_0}{\mathrm{d}z}=\frac{\mathrm{d}g}{\mathrm{d}z},\quad \frac{\mathrm{d}\alpha}{\mathrm{d}z}=\frac{1}{t}\left(\frac{\mathrm{d}g}{\mathrm{d}z}-\alpha\frac{\mathrm{d}t}{\mathrm{d}z}\right),\quad \frac{\mathrm{d}\beta_0}{\mathrm{d}z}=-\frac{\mathrm{d}g}{\mathrm{d}z},\quad \text{and}\quad \frac{\mathrm{d}\beta}{\mathrm{d}z}=-\frac{1}{t}\left(\frac{\mathrm{d}g}{\mathrm{d}z}+\beta\frac{\mathrm{d}t}{\mathrm{d}z}\right),$$

and observing also that

$$\frac{\partial Q}{\partial\alpha_0}=2\exp(-\alpha_0)+\frac{4}{3}\frac{\partial Q}{\partial\alpha},\quad \text{and}\quad \frac{\partial Q}{\partial\beta_0}=-\frac{4}{3}\frac{\partial Q}{\partial\beta},$$

Eq. (10) can be recast as

$$J_g\frac{\mathrm{d}g}{\mathrm{d}z}+J_t\frac{\mathrm{d}t}{\mathrm{d}z}=\sinh(f)Q(t,\alpha_0,\alpha,\beta_0,\beta)\left(\frac{2}{\Delta z}-\frac{\mathrm{d}f}{\mathrm{d}z}\right)+C_2J\exp(\beta_0), \tag{A12}$$

where

$$\begin{aligned}J_g&=\cosh(f)\left[Q(t,\alpha_0,\alpha,\beta_0,\beta)+2\exp(-\alpha_0)+\left(\frac{4}{3}+\frac{1}{t}\right)\frac{\partial Q}{\partial\alpha}+\left(\frac{4}{3}-\frac{1}{t}\right)\frac{\partial Q}{\partial\beta}\right],\\J_t&=\cosh(f)\left(\frac{\partial Q}{\partial t}-\frac{\alpha}{t}\frac{\partial Q}{\partial\alpha}-\frac{\beta}{t}\frac{\partial Q}{\partial\beta}\right),\end{aligned}$$

and $C_2=2\tau_0/[q\Gamma kT_0(\Delta z)^2]$.

Similarly, after neglecting the second derivative as indicated in the text, Eq. (18) can be recast making use of Eqs. (17) and (A5), thus yielding

$$H_g\frac{\mathrm{d}g}{\mathrm{d}z}+H_t\frac{\mathrm{d}t}{\mathrm{d}z}=-\exp(\beta_0)\{2C_2Jf-C_3[t^2U(t,\alpha,\beta)-U_0(\alpha',\beta')]\}-H_f\frac{\mathrm{d}f}{\mathrm{d}z}, \tag{A13}$$

with

$$\begin{aligned}H_f&=\cosh(f)S(t,\alpha_0,\alpha,\beta_0,\beta),\\H_g&=\sinh(f)\left[S(t,\alpha_0,\alpha,\beta_0,\beta)+\frac{1}{t}\left(\frac{\partial S}{\partial\alpha}-\frac{\partial S}{\partial\beta}\right)+\frac{\partial S}{\partial\alpha_0}-\frac{\partial S}{\partial\beta_0}\right],\\H_t&=\sinh(f)\left(\frac{\partial S}{\partial t}-\frac{\alpha}{t}\frac{\partial S}{\partial\alpha}-\frac{\beta}{t}\frac{\partial S}{\partial\beta}\right),\end{aligned}$$

where $C_3=\tau_0/(4\Delta z\tau_r)$.

Equations (A11)–(A13) can be turned into Eqs. (20)–(22) by letting

$$\begin{aligned}N^*&=C_1[tN(\alpha,\beta)-C_0],\\J^*&=\sinh(f)Q(t,\alpha_0,\alpha,\beta_0,\beta)\left(\frac{2}{\Delta z}-\frac{\mathrm{d}f}{\mathrm{d}z}\right)+C_2J\exp(\beta_0),\end{aligned}$$

and

$$H^*=\exp(\beta_0)\{2C_2Jf-C_3[t^2U(t,\alpha,\beta)-U_0(\alpha',\beta')]\}-H_f\frac{\mathrm{d}f}{\mathrm{d}z}.$$

[1]The International Technology Roadmap for Semiconductors, 2011, http://www.itrs.net/Links/2011ITRS/Home2011.htm.
[2]V. V. Zhirnov, R. K. Cavin III, J. A. Hutchby, and G. I. Bourianoff, Proc. IEEE **91**, 1934 (2003).
[3]D.-C. 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, and G. Spadini, Tech. Dig. - Int. Electron Devices Meet. **2009**, 617.
[4]N. Derhacobian, S. C. Hollmer, N. Gilbert, and M. N. Kozicki, Proc. IEEE **98**, 283 (2010).
[5]H.-S. Wong, S. Raoux, S. B. Kim, J. Liang, J. P. Reifenberg, B. Rajaendran, M. Asheghi, and K. E. Goodson, Proc. IEEE **98**, 2201 (2010).
[6]S. R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968).

[7]M. Wuttig and N. Yamada, Nature Mater. **6**, 824 (2007).
[8]N. Yamada, Mater. Res. Soc. Symp. Bull. **21**, 48 (1996).
[9]D. Adler, H. K. Henisch, and N. Mott, Rev. Mod. Phys. **50**, 209 (1978).
[10]D. Adler, M. S. Shur, M. Silver, and S. R. Ovshinsky, J. Appl. Phys. **51**, 3289 (1980).
[11]S. Lai and T. Lowrey, Tech. Dig. - Int. Electron Devices Meet. **2001**, 803.
[12]A. Pirovano, A. L. Lacaita, A. Benvenuti, F. Pellizzer, and R. Bez, IEEE Trans. Electron Devices **51**, 452 (2004).
[13]M. Rudan, F. Giovanardi, E. Piccinini, F. Buscemi, R. Brunetti, and C. Jacoboni, IEEE Trans. Electron Devices **58**, 4361 (2011).
[14]V. G. Karpov, Y. A. Kryukov, S. D. Savransky, and I. V. Karpov, Appl. Phys. Lett. **90**, 123504 (2007).
[15]M. Simon, N. Nardone, V. G. Karpov, and I. V. Karpov, J. Appl. Phys. **108**, 064514 (2010).
[16]D. Ielmini and Y. Zhang, J. Appl. Phys. **102**, 054517 (2007).
[17]D. Ielmini, Phys. Rev. B **78**, 035308 (2008).
[18]F. Buscemi, E. Piccinini, R. Brunetti, M. Rudan, and C. Jacoboni, J. Appl. Phys. **106**, 103706 (2009).
[19]N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon, Oxford, 1979).
[20]B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer, Berlin, 1984).
[21]C. Jacoboni, *Theory of Electron Transport in Semiconductors* (Springer, Berlin, 2010)
[22]H. Fritzsche, J. Phys. Chem. Solids **68**, 878 (2007).
[23]See supplementary material at http://dx.doi.org/10.1063/1.4761997 for an animation showing the evolution of $f(z)$, $t(z)$, and $n(z)$ as a function of the current.
[24]H. Fritzsche, IBM J. Res. Dev. **13**, 515 (1969).

Solid-State Electronics 84 (2013) 90–95

Contents lists available at SciVerse ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

Hot-electron conduction in ovonic materials

Carlo Jacoboni [a,*], Enrico Piccinini [b,*], Fabrizio Buscemi [b], Andrea Cappelli [a]

[a] Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Istituto Nanoscienze CNR-S3, Università di Modena e Reggio Emilia Via Campi 213/A, I-41125 Modena, Italy
[b] ARCES Research Center, Università di Bologna Via Toffano 2/2, I-40125 Bologna, Italy

ARTICLE INFO

Article history:
Available online 13 March 2013

Keywords:
Ovonic materials
Trap-limited conductivity
Phase-change memory

ABSTRACT

Electric conduction in ovonic materials is analyzed with special attention to chalcogenide glasses used for phase-change memories. A general theory is presented based on plausible microscopic assumptions. Electric field, carrier concentration, and electron temperature along the device, as well as diffusion and Poisson self-consistency, are considered. The effect of different ranges of localized levels in the gap is analyzed. The results account for and interpret all main experimental findings in phase-change memory cells.

1. Introduction

Ovonic materials feature an electrical switching behavior, i.e., a sudden change in resistivity by several orders of magnitude when a threshold field is reached, that cannot be explained by the standard Mott theory [1,2]. Such a behavior is of great interest for the technology of phase-change memories (PCM) [3], so that in the last decade the study of electron transport in amorphous materials raised renewed interest.

Several attempts of interpreting the S-shaped Negative Differential Resistivity (SNDR) of ovonic materials have been proposed in the literature. They rely on different microscopic pictures, ranging from the formation of filaments within an amorphous matrix [4–7] to impact ionization [8], or to space-charge accumulation [9]. Recently, transport in chalcogenide glasses was also modeled in terms of a thermally-assisted trap-limited conduction [10,11]. The electron switching is connected to a field effect that either enhances the carrier concentration in shallow traps [10], or increases the average energy of the trapped carriers [11].

These theories are shortly reviewed in Section 2.

In this work, carrier transport in ovonic materials is still investigated by means of trap-limited conduction, but following a hydrodynamic-like approach [12]. Details are reported in Section 3. Unlike the models presented in Refs. [10,11], the energy distribution of the carriers is used to calculate the local carrier concentration and their average energy, as well as the charge and energy fluxes. In the present picture, diffusion effects are explicitly included, and self-consistency between charge distribution and field is accounted for. These features were instead missing in previous works.

We also investigate the switching phenomenon as a function of the energy distribution of the trap levels. Results are reported in Section 4.

2. Available models

Several papers on transport in amorphous semiconductors refer to hopping conduction and/or to trap-limited conduction [2]. In the latter case, carriers travel in the band or in extended states above the mobility edge, but are often captured in localized trap states. Such a situation is also referred to as "reduced mobility" (see Fig. 1a) and can be described in two different ways: we may consider the total concentration of carriers n_0 with a reduced mobility μ, or the reduced concentration n of the carriers actually present in the band moving with the "free mobility" μ_0. The presence of an electric field reduces the height of the barrier to be overcome by the trapped electrons to reach the band, thus increasing the reduced mobility in one picture, or increasing the concentration of free carriers in the other picture. These two descriptions of the Poole–Frenkel effect are, of course, equivalent.

To be precise, the expression "hopping conduction" should strictly refer to the case of carriers jumping from one site to another without entering the conduction band. Such transitions may occur if two traps are close enough and the energy barrier between them is not too high. They may consist in a direct tunneling, or may be thermally assisted or thermally activated, as shown in Fig. 1b.

If τ_b is the average free time of an electron in the band before being captured (trapping time), and τ_t the average time of a trapped electron before being released to the band (escape time), the reduced mobility is easily calculated as

$$\mu = \mu_0 \frac{\tau_b}{\tau_b + \tau_t} \tag{1}$$

* Corresponding author.
E-mail addresses: carlo.jacoboni@unimore.it (C. Jacoboni), enrico.piccinini@unimore.it (E. Piccinini).

0038-1101/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.sse.2013.02.007

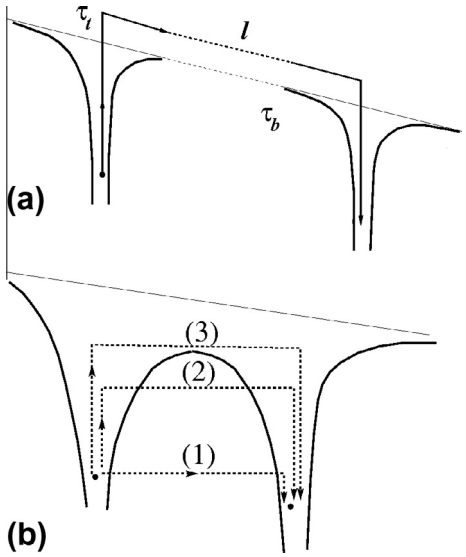

Fig. 1. (a) Reduced mobility framework. When the time τ_t spent by the carriers inside the traps is much longer than the time τ_b spent to move from one trap to another, the trap-limited mobility is characterized by two parameters: the mean flight length l and the "escape" time τ_t, as for hopping conduction. (b) Hopping may occur through a direct tunneling (1), thermally-assisted tunneling (2) or thermal activation (3). The thin line on top indicates the band mobility edge.

If $\tau_b \gg \tau_t$ electrons spend most of their time in the traps, and the above reduces to $\mu_0\,\tau_b/\tau_t$. In the last case the drift velocity becomes

$$v_d = \mu F = \mu_0 F \frac{\tau_b}{\tau_t} = \frac{l}{\tau_t}, \tag{2}$$

where l is the average distance traveled between consecutive detrapping and trapping events. Thus, only the average distance traveled by the electrons in each transition and the waiting time between successive flights are relevant, as in hopping transport.

For the ovonic materials of interest in PCM technology, the temperature and field dependences of the conductance show that transport occurs mainly via thermally-activated trap-limited mobility, with electrons moving from one site to another in the band, or above the band mobility edge [11]. It may occur that two traps are close enough to allow for direct tunneling, but a Monte Carlo simulation [13] has shown that in such a case carriers jump with high frequency back and forth between the two centers, which become equivalent to a single trap. These motions contribute significantly to the high-frequency noise, but very little to conduction.

When the research in the field of electron transport in semiconductor glasses began, two possible kinds of explanations of the switching phenomenon were proposed. One of them was based on thermal effects due to Joule heating [14]. As a matter of fact, in such materials the conductivity, being thermally activated, increases exponentially with temperature. However, this mechanism appears to be too slow to explain the experimental observations on the typical switching speed, and thermal explanations are today much less popular, even though not totally abandoned [15].

The second kind of explanations proposed, instead, purely electronic phenomena [4]. Within such a category, several mechanisms may be effective. In Ref. [16] as many as twelve mechanisms are listed and evaluated. Some of them have actually been presented in the literature without much further consideration, some have received significantly elaborations.

Here we shall limit ourselves to briefly describe those which received more attention or which are today considered of interest. Generally speaking, we must say that several of the proposed mechanisms may be present at the same time with different relative importance in various materials and conditions. Each transport theory contains a number of internal parameters that are difficult to determine. By adjusting them, it is easy to fit a single set of experimental data. Therefore, it is necessary to accumulate a large number of observations of different properties in various materials and conditions to be able to discriminate the dominant mechanism, or the right combination of the suggested mechanisms in each situation.

The ovonic switching in glassy semiconductors was initially observed by Kolomiets and Goryunova [17] and by Ovshinsky [18]. One of the first systematic works aimed at interpreting the SNDR behavior is due to Adler and coworkers, starting from the late 1970s [4,5]. The transport model they proposed relies on the simultaneous presence of electrons and holes as moving carriers, and of charged trapping centers. The interplay among these actors is such that when a threshold field is reached a complex reaction kinetics sets in. Extra pairs of carriers generated via impact ionization fill and neutralize the majority of the traps. High-conductive filaments in which carriers can move more freely are created, and the material switches into the low-resistive state.

The complex reaction kinetics and the forecast micrometric diameter of the filaments in the high-conductive state, that is no more compatible with the small cross-sections of present ovonic devices, contributed to raise a general skepticism about the Adler's interpretation. New models have then been proposed.

Still bound to the ideas of impact ionization, the model presented by the Lacaita group [8] is closer to the common framework of transport in semiconductors. Since band-to-band impact ionization seems rather impossible due to the high band gap of ovonic materials, trap states within the energy gap act here as pillars of a bridge favoring transitions between valence and conduction bands. The generation and recombination rates are not symmetric, the former being increased by larger biases, the latter being instead weakly dependent on the bias. At high currents, when generation overcomes recombination, the concentration of high-mobility carriers is increased; this implies a reduction of the bias.

Even though transitions are made easier with the help of intermediate trap states, the energy required by the carriers before ionizing is still high and does not cope with the typical mean free path of ovonic materials. However, experimental evidences suggest an activation energy for the transport process, which cannot properly be taken into account by impact ionization. An alternative interpretation for promoting carriers from the traps to the conduction band within a similar framework was recently proposed by Rudan and coworkers making use of cooperative electron–electron interaction [19].

By defining the activation energy as the energy required by a trapped electron to reach the conduction band mobility edge, Ielmini and Zhang [10] and Ielmini [11] developed a transport model in the trap-limited framework, as illustrated in Fig. 1. Under equilibrium conditions, transitions in any direction are equally probable. On the contrary, an electric field modifies the energy landscape near each trap, exponentially enhancing or limiting certain transitions. For sufficiently high fields, an almost one-directional flux of carriers is found. During their motion, carriers acquire energy from the field and release energy to the lattice by means of a relaxation process. If these energy transfers are not perfectly balanced, a heated carrier distribution is eventually formed. In turn, the heated population shows a higher mobility, which is the key condition to sustain higher currents with lower fields and to make the material switch from a low- to a high-conductive state.

Instead of interpreting the activation energy as the energy required by carriers to leave a trap, Karpov and coworkers referred to it in a thermodynamic approach [6,7]. According to their theory,

a number of crystalline nuclei are created inside an amorphous matrix through local energy dissipation. The presence of an electric field enhances the possibility to make these initial nuclei grow, until they form a filament shunting the contacts. Once the filament is created, a low-resistivity path exists in order to sustain the current, and the material can switch from a high- to a low-resistance state.

3. Hydrodynamic-like trap-limited transport model

The Ielmini's models [10,11] focus on the main physical mechanisms responsible for switching, but the use of a fixed carrier concentration is somehow contradictory with the resulting electric-field profile along the device. The model has therefore been reworked from scratch, introducing diffusion and electrostatic self-consistency [12].

If we adopt the continuum description, trap-limited conduction can be described by means of a procedure formally similar to that applied in the standard semiconductor theory where the Boltzmann transport equation holds. In particular, as in the hydrodynamic approach [20], we calculate the carrier concentration and the charge and energy fluxes by performing proper integrations of the distribution function over the energy region filled with traps. As in the standard semiconductor case, under off-equilibrium conditions a Fermi distribution function $f_F(E_F, T_e)$ incorporating a quasi-Fermi level $E_F(x, y, z)$ and a carrier temperature $T_e(x, y, z)$ replaces the equilibrium Fermi distribution.[1] For the sake of simplicity, we restrict our analysis to the one-dimensional case, where the quantities of interest are functions of the z-coordinate only, the direction of the current.

The role of the trap density of states deserves attention. In Ref. [12], a flat density of states spanning the entire band gap $\Delta E_G = E_C - E_V$ was considered, where E_C and E_V are the energies of the conduction- and valence-band edges. That choice was primarily made to keep the analytical derivation of the model simple, and it is not actually supported nor denied by experimental evidences. Even though first-principle calculations on the structure of the amorphous phase of ovonic materials are able to enlighten the defectiveness of the matrix [21–23], they still cannot provide ultimate answers about the energy distribution of the related trap states. In the lack of any further information, we speculate on the effect of the trap density of states by supposing the existence of an energy range ΔE_T around the (equilibrium) Fermi level at mid gap. In ΔE_T a flat density of states $\gamma = n_T/\Delta E_T$, where n_T is the trap concentration, is assumed.

Following Ref. [10], the average escape time of a carrier sitting in a trap at coordinate z with energy E is given by

$$\tau_{\rightleftharpoons}(E,z) = \tau_0 \exp\left[\frac{E_C - E \pm qF(z)\Delta z/2}{k_B T_L}\right], \tag{3}$$

where τ_0, q, k_B and are a characteristic time constant, the absolute value of the carrier charge, the Boltzmann constant and the lattice temperature, respectively. The arrows at the l.h.s. of Eq. (3) indicate that the escape times in the two directions of motion are different. As the escape process is supposed to be the slowest process in the conduction chain by orders of magnitude, the carrier velocity during a transition can be approximated by the ratio between the average traveling distance Δz and the escape time $\tau_{\rightleftharpoons}(E, z)$ as in Eq. (2):

$$v_{\rightleftharpoons}(E,z) = \frac{\Delta z}{\tau_{\rightleftharpoons}(E,z)} = \frac{\Delta z}{\tau_0}\exp\left[-\frac{E_C - E \mp qF(z)\Delta z/2}{k_B T_L}\right] \tag{4}$$

The above result allows for the integration of the distribution function to yield the moments that are necessary in the model:

i. carrier concentration (zero order moment)

$$n(z) = \int_{\Delta E_T} \gamma f_F(E,z)\, dE, \tag{5}$$

ii. average carrier velocities (first-order moments)

$$\langle v_{\rightleftharpoons}(z)\rangle = \frac{1}{n(z)}\int_{\Delta E_T} \gamma f_F(E,z)\, v_{\rightleftharpoons}(z)\, dE, \tag{6}$$

iii. average carrier energy (second-order moment)

$$\langle E(z)\rangle = \frac{1}{n(z)}\int_{\Delta E_T} \gamma f_F(E,z)(E - E_V)\, dE, \tag{7}$$

iv. energy fluxes (third-order moments)

$$\langle P_{\rightleftharpoons}(z)\rangle = \frac{1}{n(z)}\int_{\Delta E_T} \gamma f_F(E,z)(E - E_V)\, v_{\rightleftharpoons}(z)\, dE, \tag{8}$$

The two even moments are functions of two unknowns, the quasi-Fermi level $E_F(z)$ and the carrier temperature $T_e(z)$. The odd moments depend also on a third unknown, the electric field $F(z)$, through the carrier velocity, and can be combined together to yield transport equations. A set of three macroscopic differential equations can be written and solved for the triplet $[F(z), n(z), \langle E(z)\rangle]$ where $n(z)$ and $E(z)$ are actually written in terms of $E_F(z)$ and $T_e(z)$, as in Ref. [12]:

1. Poisson equation

$$\frac{dF(z)}{dz} = -\frac{q}{\varepsilon}[n(z) - n_{eq}], \tag{9}$$

where ε is the dielectric constant of the material and n_{eq} is the equilibrium carrier concentration. Under equilibrium conditions, the system is neutral at any z.

2. Carrier flux equation

$$\frac{J}{q} = -n(z)[\langle v_{\rightarrow}(z)\rangle - \langle v_{\leftarrow}(z)\rangle] + \frac{\Delta z}{2}\frac{d}{dz}\{n(z)[\langle v_{\rightarrow}(z) + v_{\leftarrow}(z)\rangle]\}, \tag{10}$$

Since the model must be current-driven, as required for a proper investigation of typical S-shaped $I(V)$ characteristics, the l.h.s. is known *a priori*. The r.h.s. consists of two terms. The former stems from drift, whereas the latter takes into account the non-uniformity of the carrier concentration and velocity, and represents the diffusive contribution to the current.

3. Energy balance equation

$$\frac{d}{dz}\{n(z)[\langle P_{\rightarrow}(z)\rangle - \langle P_{\leftarrow}(z)\rangle]\} - \frac{\Delta z}{2}\frac{d^2}{dz^2}\{n(z)[\langle P_{\rightarrow}(z)\rangle + \langle P_{\leftarrow}(z)\rangle]\} = JF(z) - \frac{n(z)[\langle E(z)\rangle - \langle E_0(z)\rangle]}{\tau_R} \tag{11}$$

This equation has been derived in the first-order approximation of the energy relaxation through a relaxation time constant τ_R and a "target energy" $\langle E_0(z)\rangle$, namely the energy that the population $n(z)$ would have if at thermal equilibrium with the lattice at temperature T_L.

The analytical derivation of these equations is reported in Ref. [12] for the case of traps filling the entire band gap.

4. Results

The results obtained from the theory presented in the previous section are shown in Figs. 2–6. Fig. 2 displays the $I(V)$ curves exhibiting the standard features of the conduction in ovonic materials, i.e., an Ohmic region at the lowest current, followed by an

[1] In fact, a very accurate approximation of the Fermi distribution that makes the analytical integration possible has been used, as in Ref. [12].

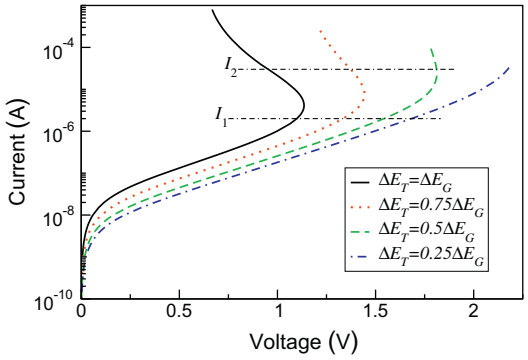

Fig. 2. $I(V)$ characteristics of a 40 nm-long device with a 1000 nm^2 cross-sectional area for different values of ΔE_T ranging from ΔE_G down to $0.25\Delta E_G$, with $\Delta E_G = 0.68$ eV. The parameters used in the numerical calculations are $\Delta z = 7$ nm, $\tau_0 = 1.2 \times 10^{-14}$ s, $\tau_R = 7.8 \times 10^{-14}$ s, $n_T = 6.8 \times 10^{19}$ cm^{-3} and $T_0 = 300$ K. The values of these parameters are the ones used in Ref. [12]. The two horizontal lines indicate the two currents $I_1 = 2 \times 10^{-6}$ A, and $I_2 = 3 \times 10^{-5}$ A for which the profiles of $F(z)$, $T_e(z)$, and $n(z)$ are shown in Fig. 4.

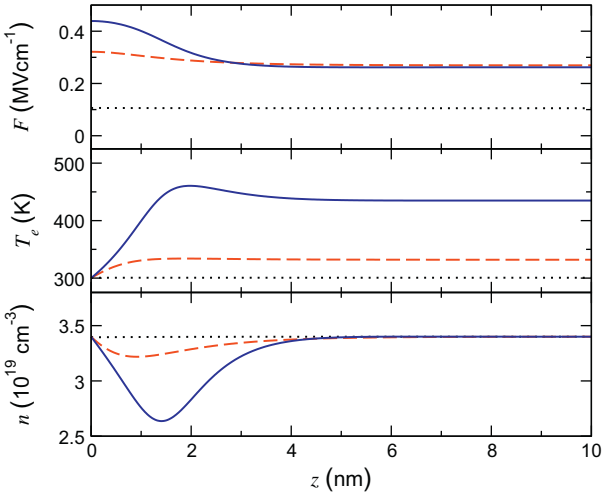

Fig. 3. Profiles of $F(z)$ (top panel), $T_e(z)$ (middle panel), and $n(z)$ (bottom panel) in the region of the device from 0 to 10 nm for the solid (black) curve of Fig. 2 evaluated at three selected currents: 1×10^{-7} A (dotted lines), 2×10^{-6} A (dashed lines), and 1×10^{-5} A (solid lines).

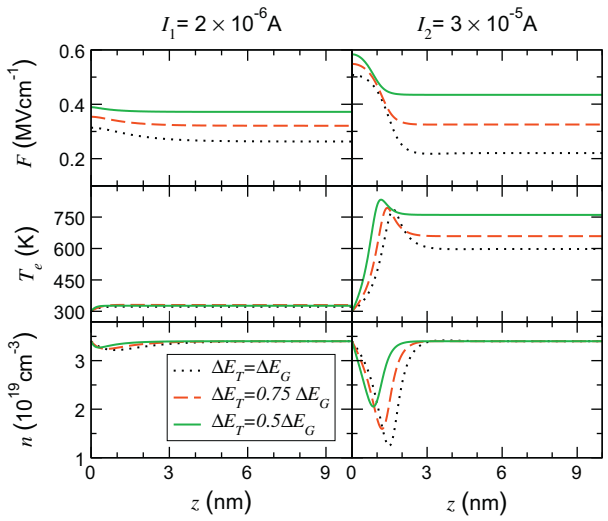

Fig. 4. Profiles of $F(z)$ (top panels), $T_e(z)$ (middle panels), and $n(z)$ (bottom panels), in the region of the device from 0 to 10 nm, when $\Delta E_T = \Delta E_G$ (dotted line), $\Delta E_T = 0.75\ \Delta E_G$ (dashed line), and $\Delta E_T = 0.5\ \Delta E_G$ (solid line), as in Fig. 2. For longer z's all curves keep their saturation values.

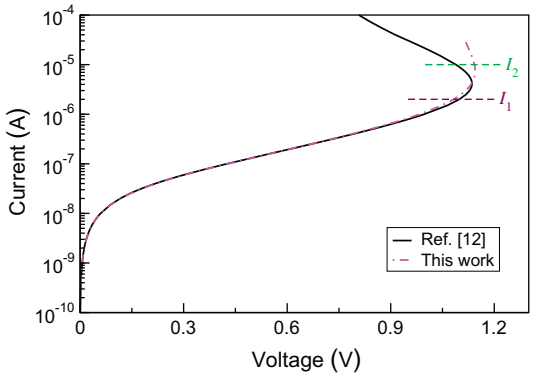

Fig. 5. Comparison between the curve of Ref. [12] reproducing the best fit of experimental data and the curve here obtained with $\Delta E_T = 0.5\ \Delta E_G$, $\tau_0 = 4.2 \times 10^{-15}$ s, and $\tau_R = 4.2 \times 10^{-13}$ s. The dashed horizontal lines show two significant currents that are analyzed in Fig. 6.

exponential regime. When the electric field reaches a threshold value, a voltage snap-back occurs.

In the subthreshold region, the profiles of the electric field and of the carrier temperature are essentially uniform, the latter being also close to the lattice temperature (see Fig. 3). As a consequence, the carrier concentration is constant. Any increase in the input current leads simply to a spatially-uniform growth of the electric field.

The situation dramatically changes when the threshold condition is approached. The spatial uniformity of the three profiles is attained only after a region close to the injecting contact, where the quantities of interest vary appreciably. What happens within this region is responsible for the switching effect. Due to the high electric field, a high power is also transferred to the carriers. If the energy-relaxation mechanism is no more able to transfer completely that power to the lattice, carriers tend to heat up while moving from the injecting to the collecting contact. As their average energy (thus temperature) increases, carriers move faster and, in order to keep the current equal to the prescribed value, the carrier concentration $n(z)$ decreases. In turn, hotter carriers can overcome energy barriers more easily, and a lower field is required to sustain the current. Moreover, lower fields also mean a lower input power. For a sufficiently high carrier temperature, the energy relaxation balances again the energy gain from the field. From that point on, the spatial uniformity of the field, of the carrier temperature, and of the carrier concentration is also restored. This "on" region is much longer than the first one, close to the injecting contact, where carriers are heated up ("off" region). The effect of the field in this "off" region on the overall potential drop is extremely reduced. The value of the field in the "on" region can be assumed as almost representative of the whole device. While the field close to the injecting contact always increases with the current, the field in the "on" region, thanks to the interplay between energy gain and energy relaxation, increases at a slower rate in the last part of the subthreshold region and eventually decreases, giving origin to the threshold-switching effect.

In the following, we will focus on the influence played by the energy range ΔE_T of the trap levels. To this aim, in Fig. 2 four $I(V)$ characteristics, obtained with a progressively narrowed ΔE_T at a fixed trap space concentration n_T, are displayed. The results of the calculations clearly indicate that the conductivity in the subthreshold region decreases and the threshold voltage and current increase.

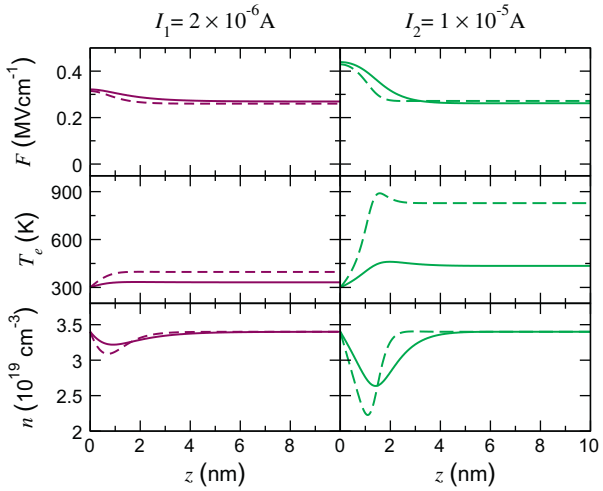

Fig. 6. Profiles of $F(z)$, $T_e(z)$ (middle panels), and $n(z)$ (bottom panels) for the two selected currents, namely 2×10^6 A and 1×10^5 A, of the $I(V)$ characteristics of Fig. 5. Solid lines refer to the curve of Ref. [12], while the dashed lines refer to the curve obtained in this work. The z axis has been truncated at 10 nm to highlight the region close to injecting contact where $F(z)$, $T_e(z)$ vary most appreciably.

The lower conductivities found for smaller energy ranges can be attributed to the suppression of the shallow states within the band gap. According to Eqs. (3) and (4), if the trap energies are farther away from the bottom of the conduction band, the escape time is longer, and the carrier velocity lower. On the opposite side, the suppression of the very deep states does not alter significantly the results, as their contribution to transport is substantially negligible.

In Fig. 4 the profiles of the electric field $F(z)$, carrier temperature $T_e(z)$ and carrier concentration $n(z)$ are reported as a function of the position along the device for two currents around the threshold value. Note that the threshold point moves roughly from $I_1 = 2 \times 10^{-6}$ A to $I_2 = 3 \times 10^{-5}$ A when the range of the trap energy levels is reduced by a factor of 2 (solid line vs. dashed line in Fig. 2).

The lower conductivities obtained for reduced energy ranges imply that a given current is sustained only if the energy barriers between traps are further reduced by the action of a higher electric field. Since the field is almost constant along the device but for a narrow region close to the injecting contact, higher overall potential drops are found, as can be seen in Figs. 2 and 4. In the very first part of the device, the average energy of the carriers is increased by the imbalance between energy gain and relaxation. Having suppressed the shallow energy levels in the band gap, the carrier average energy necessary for the switching event is obtained only in presence of higher fields and currents able to strongly increase the temperature of the carriers whose energies are close to the Fermi level, as reported in Fig. 4. As a consequence, larger threshold currents and voltages occur.

Despite the variation in the energy range of the traps, it is always possible to find reasonable pairs of the characteristic times τ_0 and τ_R to reproduce experimental data. In particular, small ranges require shorter characteristic times τ_0 and longer relaxation times τ_R. This result is consistent with the previous considerations about the suppression of shallow energy levels. In Fig. 5, we show two $I(V)$ characteristics obtained when $\Delta E_T = \Delta E_G$ ($\tau_0 = 1.2 \times 10^{-14}$ s and $\tau_R = 7.8 \times 10^{-14}$ s) [12] and when $\Delta E_T = 0.5\Delta EG$ ($\tau_0 = 4.2 \times 10^{-15}$ s and $\tau_R = 4.2 \times 10^{-13}$ s). The two curves are almost identical up to the threshold point where any difference between them is well within the experimental errors. A discrepancy is present above the threshold point and can be attributed to the different locations of the quasi-Fermi levels that may reach or trespass the lower edge of ΔE_T in the region close to the contact. This situation represents a natural limit of the present model. However, the availability of experimental results not affected by the external circuitry would provide an important source of information to discriminate the range ΔE_T.

We have pointed out that near the switching conditions the profile of the carrier temperature depends on the trap energy range. In Fig. 6, we report the profiles of $F(z)$, $T_e(z)$, and $n(z)$ for the two conditions $\Delta E_T = \Delta E_G$ (solid lines) and $\Delta E_T = 0.5\Delta E_G$ (dashed lines). As the input power density is $JF(z)$, one notes that the power provided by the electric field is roughly the same for the two conditions considered, and the same must hold true also for the energy gain in the region close to the injecting contact. The effect of suppressing the shallow energy levels implies a larger increase of the temperature of the carriers around the Fermi level. This effect introduced by narrow distributions and the corresponding change in the value of τ_0 necessary to maintain the agreement with the experimental $I(V)$ curve, make the carriers move faster. As a consequence, the "off" region is more depleted with respect to larger trap-energy ranges. From the Poisson equation, it follows that the electric field must vary more steeply. Due to these steeper variations, the main difference between the two conditions considered is the spatial extension of the "off" region, that becomes progressively shorter as ΔE_T narrows.

5. Conclusions

In this paper, a theory of electron transport in ovonic materials, based on a very general and plausible physical model, is predented. The current is produced by carriers with trap-limited mobility. Experimental data are well reproduced below and around the switching point. The voltage snap-back is produced by the heating of the carriers, as in Refs. [11,12]: in order to sustain a high current a large field is necessary near the injecting contact to lower the barrier to be overcome by the carriers. These are therefore heated up and acquire a large mobility, whose value is maintained in the second and longer part of the device. Above threshold, a larger current requires a larger initial field which produces larger carrier heating and mobility so that the applied voltage is reduced. According to the present model, trap energy levels are distributed with constant density of states in a range ΔE_T in the middle of the band gap.

Particular attention has been paid to the influence of the amplitude of ΔE_T on the transport properties of the material. When such a range is narrowed, the overall conductivity decreases as the carriers must overcome a higher barrier to reach the conduction band. The switching point is moved to higher voltages and currents. The extension of the high-field region near the injecting contact is reduced with possible implications of great practical interest in device technology.

Acknowledgments

The authors would like to thank Rossella Brunetti, Fabio Giovanardi, Daniele Ielmini, Massimo Rudan, and Gianpaolo Spadini for many useful discussions.

We gratefully acknowledge the Intel Corporation support under the contract 3477131/2011.

References

[1] Mott N, Davis E. Electronic processes in noncrystalline material. Oxford: Clarendon Press; 1961.
[2] Shklovskii B, Efros AL. Electronic properties of doped semiconductors. Berlin: Springer; 1984.
[3] Kau D et al. A stackable cross point phase change memory. In: 2009 IEEE international electron devices meeting (IEDM); 2009. p. 1–4.

[4] Adler D, Henisch HK, Mott SN. The mechanism of threshold switching in amorphous alloys. Rev Mod Phys 1978;50:209.
[5] Adler D, Shur MS, Silver M, Ovshinsky SR. Threshold switching in chalcogenide-glass thin films. J Appl Phys 1980;51:3289.
[6] Karpov VG, Kryukov YA, Savransky SD, Karpov IV. Nucleation switching in phase change memory. Appl Phys Lett 2007;90:123504.
[7] Simon M, Nardone M, Karpov VG, Karpov IV. Conductive path formation in glasses of phase change memory. J Appl Phys 2010;108:064514.
[8] Pirovano A, Lacaita A, Benvenuti A, Pellizzer F, Bez R. Electronic switching in phase-change memories. IEEE Trans Electron Dev 2004;51:452.
[9] Buscemi F, Piccinini E, Brunetti R, Rudan M, Jacoboni C. Monte Carlo simulation of charge transport in amorphous chalcogenides. J Appl Phys 2009;106:103706.
[10] Ielmini D, Zhang Y. Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices. J Appl Phys 2007;102:054517.
[11] Ielmini D. Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses. Phys Rev B 2008;78:035308.
[12] Piccinini E, Cappelli A, Buscemi F, Brunetti R, Ielmini D, Rudan M, et al. Hot-carrier trap-limited transport in switching chalcogenides. J Appl Phys 2012;112:083722.
[13] Buscemi F, Piccinini E, Rudan M, Brunetti R, Jacoboni C. Diffusion and high-frequency noise of electrons in amorphous semiconductors at low electric fields. Fluct Noise Lett 2012;11:1242004.
[14] Owen A, Robertson J, Main C. The threshold characteristics of chalcogenide-glass memory switches. J Non-Cryst Solids 1979;32:29.
[15] Bogoslovskiy N, Tsendin K. Physics of switching and memory effects in chalcogenide glassy semiconductors. Semiconductors 2012;46:559.
[16] Nardone M, Simon M, Karpov IV, Karpov VG. Electrical conduction in chalcogenide glasses of phase change memory. J Appl Phys 2012;112:071101.
[17] Kolomiets B, Goryunova NA. Svoistva i struktura troinykh poluprovodnikovykh sistem. 1. Elektricheskie svoistva i struktura nekotorykh materialov v sisteme Tl–Sb–Se. Zh Tekh Fiz 1955;25:984.
[18] Ovshinsky SR. Reversible electrical switching phenomena in disordered structures. Phys Rev Lett 1968;21:1450–3.
[19] Rudan M, Giovanardi F, Piccinini E, Buscemi F, Brunetti R, Jacoboni C. Voltage snapback in amorphous-gst memory devices: transport model and validation. IEEE Trans Electron Dev 2011;58:4361.
[20] Taschini S, Rudan M, Brunetti R. Particle and energy fluxes in semiconductors: full-band hydrodynamic equations and the thermodynamic limit. Phys Rev B 1999;60:13582–91.
[21] Akola J, Jones RO. Structural phase transitions on the nanoscale: the crucial pattern in the phase-change materials $Ge_2Sb_2Te_5$ and GeTe. Phys Rev B 2007;76:235201.
[22] Caravati S, Bernasconi M, Khne TD, Krack M, Parrinello M. First-principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects. J Phys: Condens Matter 2009;21:255501.
[23] Cho E, Im J, Park C, Son WJ, Kim DH, Horii H, et al. Atomic and electronic structures of amorphous $Ge_2Sb_2Te_5$; melt-quenched vs. ideal glasses. J Phys: Condens Matter 2010;22:205504.

4.1 Summary

In this Chapter, a quasi-1D model for charge transport in Ovonic materials was presented. The basic assumptions are trap-limited conduction, and carrier energy gain and loss, due to acceleration by an electric field and interaction with phonons. These concepts are coupled with electrostatic self-consistency of the system, through the Poisson equation. This represents an improvement upon the previous literature, since the resulting variable carrier concentration is coherent with the electric field distribution, and it is employed to properly calculate the diffusive contributions to the current (including non-uniformities of carrier concentration and local carrier temperature).

Experimental data for the subthreshold and around-switching regions are nicely fitted, while no prediction is made about the IV curve well-above the voltage snap-back, since it depends on external circuit features [40], which are not included in the present model. The observed dependencies on device parameters, such as threshold voltage versus device length, or switching conditions dependence on temperature [39], are correctly recovered. Also, the effect of potentially different trap state distributions in the bandgap is assessed by assuming a uniform density of states around midgap, with variable width. It is found that a narrower energy width of the localized-state region hinders the conductivity of the material, since electrons must typically be excited from deeper states and consequently experience a higher energy barrier for conduction; the same conditions also cause an increase of the electric field required for switching.

Some limitations are still present, though: the model is the 1D projection of a 3D device. Furthermore, transport is tackled in a continuum approximation, where position and energy uniformity can average out the potential effects of discreteness (traps) and/or randomness (disorder) of the localized states in an amorphous sample. Further improvement along these lines are described in Chapter 5.

Chapter 5

Random Network Model

In Chapter 4 we described a trap-limited quasi-1D model for charge transport in Ovonic switching devices, assuming the existence of a continuity of states in both position and energy. It means that localized states are available in a continuous range of energies inside the gap and at all points along the device length. This has some similarities to treating the bandgap itself as an effective special conduction band, with transport obeying its own Poole-like law.
This, however, inevitably neglects the effect of disorder, since uniformity makes macroscopically identical devices equal to each other also microscopically, which is not the case for amorphous samples. Furthermore, localized states can be better modeled as discrete dots. This can be relevant for the chalcogenide case [85, 128, 81, 54, 125] as well as for many other devices exhibiting switching between low- and high-conductivity states [134, 124, 113, 109], the former often due to an amorphous or defective component.

These considerations suggested to adopt a new approach, based on a network of randomly generated sites, which represent the localized states in a disordered medium.
Among the possible physical pictures (e.g., [45, 123, 79], see also Chapters 3 and 4), we assumed the Poole-like emission-travel-recapture concept described in [48, 47, 91, 50, 19], typically enabled by interactions with phonons [76], since this picture complies with several experimental evidences [48] and was proved able to capture the switching behavior. Each network site hosts a population of carriers that can move among nodes according to the above-mentioned Poole process, mimicking trap-limited conduction with trap-to-trap transition rates. We also introduce a simplifying assumption such that the interaction among nodes is limited within a cutoff distance r_{cut}, and farther nodes do not exchange particles. Electrons in the nodes are also characterized by an average energy, so that

moving carriers imply an energy flow among nodes, modified in the model according to the energy gained by the electric field and the energy relaxed via phonon emission. Electrostatics is considered through the Poisson equation, that relates the electric potential in all points of the device to the population in the localization sites, and is solved via a Finite Element Method (FEM) algorithm.
This framework allows to write steady-state balance equations for both particle- and energy-fluxes at each node, which, coupled with the Poisson equation, finally allow one to find the voltage drop at the device ends, for each prescribed input current.

The random nature of the model (i.e. the network generation at the beginning of the computational procedure) allows us to study the effect of disorder in a statistical sense, since nominally identical nanometric devices can show somewhat different electrical properties due to their different site configurations. This feature of the model can be exploited to assess the typical device variability as a function of geometry and material properties, particularly as regards key technological parameters such as the (V_{th}, I_{th}) OTS switching point.
The simulated IV curves are then compared to data from GST memory cells contacted using carbon nanotubes (CNT-GST) [130, 129].
Finally, the microscopic situation of the carriers in the random network is analyzed [97], by observing the carrier energies in the nodes and the particle fluxes among them, for several points of the IV curve [3], and evidencing the formation of preferential paths around the threshold point.
We will present this approach in what follows, reproducing the original paper [27].

Conductive preferential paths of hot carriers in amorphous phase-change materials

Andrea Cappelli,[1] Enrico Piccinini,[2,a)] Feng Xiong,[3] Ashkan Behnam,[3] Rossella Brunetti,[1] Massimo Rudan,[2] Eric Pop,[3,b)] and Carlo Jacoboni[1]
[1]*Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Via Campi 213/A, I-41125 Modena, Italy*
[2]*"E. De Castro" Advanced Research Center on Electronic Systems (ARCES), Università di Bologna, Via Toffano 2/2, I-40125 Bologna, Italy*
[3]*Department of Electrical and Computer Engineering, Micro and Nanotechnology Lab, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

(Received 28 June 2013; accepted 2 August 2013; published online 20 August 2013)

We study charge transport properties of amorphous phase-change materials (PCM) using a set of balance equations applied to a three-dimensional random network of sites. In the context of trap-limited conduction, model results are checked against experimental data on PCM devices near the limits of scaling (~10 nm), explaining the main features of the current-voltage characteristics. The stochastic nature of the network also allows us to investigate the statistical variability of the sub-threshold PCM operation. Simulations of batches of similar samples show a standard deviation for the threshold condition of the order of few percent for the threshold voltage and of ten percent for the threshold current. The analysis of the network at the microscopic level near threshold reveals the formation of high-current paths, connecting the two contacts of the device through network nodes hosting the hottest carriers. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4819097]

Understanding the transport properties of disordered materials is of great interest from both the fundamental and the applied points of view. At present, the electronic industry is manufacturing an increasing number of devices whose operation relies on the properties of amorphous materials. Within the field of data storage, materials like chalcogenides,[1,2] silicon dioxide,[3] metal oxides,[4] graphene,[5] and amorphous carbon[6] can controllably be switched between a low- and a high-resistance state. In many cases, the low conductivity state stems from a disordered and defective structure. These properties allow for reversible information encoding and its readout.[7–9]

Such amorphous materials often show an onset of so-called threshold-switching phenomena when proper conditions are reached. The microscopic knowledge of the physical processes involved is of the utmost importance for any technological exploitation. In the particular case of the Ge-Sb-Te (GST) chalcogenide compounds, electronic switching leads to a large increase in conductivity triggered by the application of a voltage above a threshold value V_{th}.[9]

Although thermal interpretations for Ovonic switching have not been completely abandoned,[10] the most often referenced models ascribe this threshold-switching effect to electronic phenomena, such as hot-electron trap-limited transport[11–13] or field-induced crystallization.[14] Due to the complexity of the phenomenon, these theories usually introduce a quasi one-dimensional (1D) approximation or consider a three-dimensional (3D) continuous medium, averaging out the intrinsic disorder of the amorphous material. With the miniaturization of nanoelectronics, very small numbers of carriers and traps are involved in the device operation. Rough estimates indicate that only about one hundred defect states are present in phase-change memories at the 32 nm technological node. Thus, a more realistic 3D description of the positions of the defects paves the way to better and more precise modeling.

In this work we describe amorphous GST by means of the nodes of a random network. The node sites represent the actual positions of the traps of the real system, or the average coordinates of clusters of neighboring traps. Details on the positions of the traps should be provided by specific experimental investigations or by first-principle calculations that are not available within the required precision. Therefore, positions are randomly generated within a 3D domain according to uniform distributions along each coordinate. A minimum distance r_{min}, representing the typical dimension of the clusters, separates any pair of nodes. The transport picture complies with the hydrodynamic-like (HD) model for trap-limited conduction[12,13] and assumes carriers sitting at the node sites and moving among them according to an emission-travel-recapture process. For a prescribed input current flowing across the device, we find the corresponding voltage drop at its ends, calculating a characteristic current-voltage (I–V) curve. The effect of different node configurations on the conduction behavior and on the switching point is also assessed. The calculated I–V curves are compared to experimental data from carbon-nanotube-contacted GST-based (CNT-GST) devices.[15,16] Finally, by considering the carrier flow among the network nodes, we observe the appearance of preferential hot-carrier conduction paths, occurring in proximity of the electrical switching.

We consider a device made of an amorphous GST layer sandwiched between two conductive surfaces and model it

a)Author to whom correspondence should be addressed. Electronic mail: enrico.piccinini@unimore.it
b)Present address: Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA

as a 3D network of N randomly placed nodes. At two opposite edges, two special planar equipotential nodes (0 and $N+1$) play the role of the contacts. We associate to the i-th node its electrostatic potential φ_i and the carrier population n_i. The average energy ε_i is attributed to all members of the population sitting at node i. The zero of ε_i corresponds to the equilibrium distribution.

Following the concept of trap-limited conduction and considering for simplicity only the case of electrons, carrier motion comprises three stages: emission of a carrier from the i-th node to a traveling state with energy just above the conduction-band mobility edge E_C, propagation toward the j-th node, and finally recapture by the j-th node. To represent the overall process we assume the transition rate

$$S_{ij} = \frac{f(r_{ij})}{\tau_0} \exp\left(-\frac{E_C - \varepsilon_i}{k_B T}\right) \exp\left[\frac{-q(\varphi_i - \varphi_j)\ell}{r_{ij} k_B T}\right], \quad (1)$$

where $\tau_0 \sim 10-100$ fs is the characteristic time of the transition process (dominated by the detrapping process, most likely due to the interaction with phonons),[17] k_B is the Boltzmann constant, T the lattice temperature, q the elementary charge, and r_{ij} the distance between nodes i and j. The term $f(r_{ij})$ in Eq. (1) is a function of the internode distance r_{ij}, describing propagation and recapture. For simplicity we take $f(r_{ij}) = 0$ for r_{ij} larger than a cutoff distance r_{cut}, $f(r_{ij}) = 1$ otherwise. Furthermore, the two exponentials account for the thermally activated nature of the excitation process.[18,19] The first one accounts for the energy barrier in the absence of field; the second one is an effective correction given by the average electric field $F = -(\varphi_j - \varphi_i)/r_{ij}$ between nodes i and j at a characteristic distance $\ell \lesssim r_{cut}$. When $r_{ij} < 2\ell$, the correction is evaluated at $\ell = r_{ij}/2$, similarly to Ref. 20.

The connections among the nodes define a network of transition rates; given any pair of nodes (i, j), the current between them is $I_{ij} = -q(n_i S_{ij} - n_j S_{ji})$. In order to find the steady-state value of the node variables $(\varphi_i, \varepsilon_i, n_i)$, we impose charge- and energy-flux balances at each node. As for the latter, three contributions are taken into account: the energy lost at the node site due to carriers moving away, the energy gained from the carriers arriving at the node corrected by the effect of the field experienced during the transition, and the energy transferred to the lattice via phonon scattering with a relaxation time τ_R. Neglecting higher order-terms, for a given current I, the two balance conditions[21] yield

$$\frac{I}{(-q)}\delta_{0,i} + \sum_j n_i S_{ij} = \sum_j n_j S_{ji} + \frac{I}{(-q)}\delta_{N+1,i}, \quad (2)$$

$$\sum_j n_j S_{ji}[\varepsilon_j - q(\varphi_j - \varphi_i)] + \frac{I}{(-q)}\varepsilon_i \delta_{0,i} = \varepsilon_i \left[\sum_j n_i S_{ij} + \frac{I}{(-q)}\delta_{N+1,i}\right] + n_i \frac{\varepsilon_i - \varepsilon_{i,eq}}{\tau_R}, \quad (3)$$

with $i = 0, 1, \ldots, N+1$. Since the model is current driven, the two equations for nodes 0 and $N+1$ carry two Kronecker's δ extra terms that prescribe the fluxes from and to the external circuit. The two sets of non-linear equations above are solved along with the Poisson equation to find φ_i, ε_i, and n_i in all the device nodes, employing a Newton-Raphson scheme coupled to a finite-element procedure. Given the network configuration, the algorithm leads to an I–V curve, with $V = \varphi_{N+1} - \varphi_0$.

The configuration-dependent nature of the model allows one to study the impact of intrinsic disorder on the threshold conditions. Figure 1 shows a set of one hundred I–V curves, differing only in the particular random node positions. The statistical analysis of the different threshold voltages and currents shows variations of up to ~15% for V_{th} and ~40% for I_{th}, with respect to the corresponding average values $\langle V_{th}\rangle$ and $\langle I_{th}\rangle$. Despite quantitative differences, the three known conduction regimes identified by the analytical models (Refs. 11 and 12)—Ohmic, exponential, and super-exponential—are recovered in all curves, together with the onset of the voltage snap-back. As in Ref. 12, the existence of two stable states for a given voltage is associated with the realization of two microscopic configurations, one with a near-equilibrium carrier mean energy, the other with a non-equilibrium hot-carrier distribution.

To validate our model against experimental data, we consider a theoretical I–V curve obtained for a given set of parameters averaging the potential drops of 30 simulations with random configurations. The standard deviation of the potential drops provides an estimate of the expected variability. In Fig. 2, we show the calculated average curves and standard deviations best fitting three experimental memory device I–V curves. The experimental data are taken from recent GST devices with CNT electrodes, which represent the smallest GST volumes probed to date. The CNT electrodes are ~2 nm in diameter, enabling probing of ~10 nm scale GST bits, which switch with sub-1 μA currents, nearly 100× lower than previous state-of-the-art.[16] Valuable models for similar devices would therefore enhance the

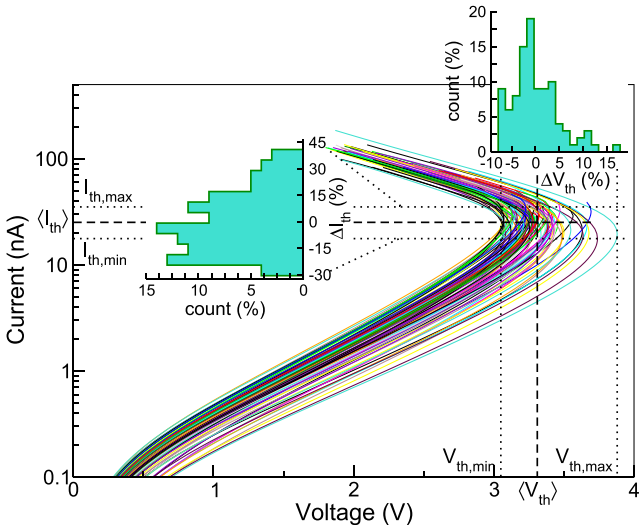

FIG. 1. Statistical fluctuations in the I–V curves for different spatial configurations of the network. The dashed and the dotted lines show the average and the minimum and maximum values for V_{th} and I_{th}, respectively. The two insets report the frequency analysis for the occurrence of the deviations from the average values $\langle V_{th}\rangle$ and $\langle I_{th}\rangle$. The analysis has been performed on a test device of $10 \times 10 \times 40\,\text{nm}^3$ setting the following parameters: $\tau_0 = 10^2$ fs, $\tau_R = 10^3$ fs, $r_{cut} = 6$ nm, $\ell = 2$ nm, node concentration $1.2 \times 10^{19}\,\text{cm}^{-3}$.

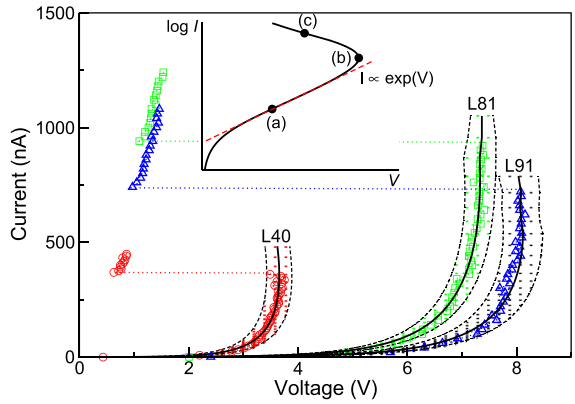

FIG. 2. Experimental *I–V* characteristics of three different CNT-GST devices (see Ref. 16), whose lengths are 40 nm (red circles), 81 nm (green squares), and 91 nm (blue triangles). The estimation of the device size is provided in the supplementary material.[21] The black solid lines and the dashed side lines show, up to the threshold region, the simulated average characteristics and the standard deviations obtained running a set of 30 simulations in each case. The inset shows a typical *I–V* characteristic in a semilog scale; the dashed lines are intended as an eye-guide for the exponential regime. Points (a), (b), and (c) used in the microscopic analysis of Fig. 3 are also highlighted.

technological development of next-generation non-volatile memories. In the subthreshold region the agreement is good. The experimental points lie within the dispersion range, confirming that the fitted parameters are reasonable. The onset of the threshold switching in the theoretical curve, i.e., the point where $dV/dI \approx 0$, corresponds quite accurately to the experimental threshold point. The abrupt voltage drop of the experimental curves above threshold eventually involves the crystallization of the system,[9] which is not addressed by the model in its present form.

The parameters used for the fit are listed in Table I and compare fairly well to those obtained with the HD model for GST memory devices.[12,21] Parameters τ_0, r_{cut}, and ℓ are similar in the three cases, so that they can be taken as best fit values of material properties. A larger variability is instead found for parameter τ_R. On the other hand, we notice that the variability may also depend on the degree of amorphization and on geometric effects of the sample that have not been considered here. The values obtained for the characteristic time τ_0 and the relaxation time τ_R confirm the link of these parameters with phononic interactions.[17] The connection of ℓ and r_{cut} with the mean propagation distance is detailed in the supplementary material.[21]

The possibility of tracking conductive preferential paths is one of the features of the present model. Figure 3 shows three frames of a ball-and-stick animation[21] from a single network configuration, representing the mean energies ε_i on the nodes and the normalized currents I_{ij}/I, for the three points (a), (b), and (c) indicated in the inset of Fig. 2. From the comparative analysis of the images, one easily sees how the different transport behaviors originate from different microscopic conditions. Figure 3(a) refers to point (a) in Fig. 2 and is obtained for a current in the exponential regime, sufficiently far from the threshold point. In this case, the current is almost evenly distributed among several paths, each of them carrying a similar fraction of the total current, roughly inversely proportional to the number of links crossing a section. Slight non-uniformities in the flux distribution are generated by geometrical factors due to the particular positions of the nodes. The carrier energies are all close to the equilibrium value. As the switching point is approached, the picture rapidly changes [Fig. 3(b)]: the non-uniformities become more evident, and the current flowing through the geometrically favored nodes is enough to induce heating of their carrier populations. This, in turn, enhances their transition rates and allows an even larger current to flow on those particular paths. The outcome of the process is the positive feedback needed to trigger the switching event: more current implies more heating and faster transition rates, and so on. Such a loop determines a steady-state condition where one or a few highly conductive continuous routes connect the contacts

TABLE I. Model parameters for highly scaled GST memory devices with CNT electrodes, as from the three best fits of Fig. 2. A cross-section $10 \times 10\,\text{nm}^2$ and a node concentration $1.2 \times 10^{19}\,\text{cm}^{-3}$ have been adopted in all cases.

Device length (nm)	τ_0 (fs)	τ_R (fs)	r_{cut} (nm)	ℓ/r_{cut}
$L = 40$	49	56.5	6	0.48
$L = 81$	49.5	25	6.5	0.5
$L = 91$	58.3	33.3	6.5	0.5

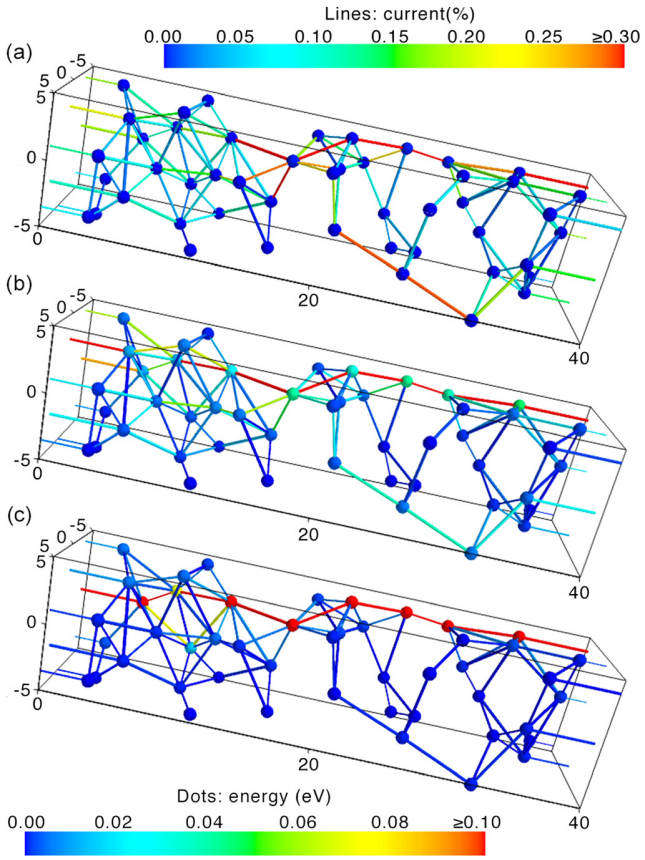

FIG. 3. Snapshots of the microscopic network configurations (lines – top scale: current; dots – bottom scale: node energy) obtained for the three points indicated in Fig. 2 for a $10 \times 10 \times 40\,\text{nm}^3$ device. The dynamics of formation of a preferential path is evidenced by: (i) the increasing number of the low-resistance connections; (ii) the progressive reduction of the current flowing in the rest of the device; and (iii) the growing energy of the nodes along the preferential path. Images have been created with the visualization software Mayavi (see Ref. 22).

TABLE II. Effect of the interaction length r_{cut} on the threshold point and its variability. Mean values and standard deviations are reported for both V_{th} and I_{th} for groups of 100 simulated devices.

r_{cut} (nm)	$\langle V_{th}\rangle \pm \sigma_{V_{th}}$ (V)	$\langle I_{th}\rangle \pm \sigma_{I_{th}}$ (nA)
6	3.31 ± 0.16 (±4.8%)	25.17 ± 4.35 (±17.3%)
7	2.94 ± 0.07 (±2.4%)	33.71 ± 4.48 (±13.3%)
8	2.74 ± 0.05 (±1.8%)	40.90 ± 5.71 (±14.0%)
10	2.48 ± 0.04 (±1.6%)	52.50 ± 6.86 (±13.1%)
12	2.29 ± 0.03 (±1.3%)	65.21 ± 8.19 (±12.5%)

[Fig. 3(c)], carrying most of the prescribed current. The lower resistance of these paths leads to a decrease in the voltage dropped across the device.

We remark in passing that the power dissipated along these preferential paths could be enough to trigger crystallization near the threshold condition. In order to further confirm this hypothesis, the present model must be enhanced with the incorporation of the Fourier heat equation and a local lattice temperature. Such an extension would also help in describing the abrupt voltage drop of the experimental *I–V* characteristics of Fig. 2.

The concept of filamentation was introduced by Adler and coworkers[23] for large samples. However, the sizes of filaments predicted in this classical case do not appear compatible with nanoscale devices.[16] Although the effect on the *I–V* characteristic is similar, the preferential paths identified by our model have a different origin than the filaments of Adler's model: in the present case, they are associated with non-equilibrium energy transport, and not with a change in the quality or charge of the traps. The paths just indicate preferred sequences of connected nodes without any reference to the filament size.

This microscopic picture complies with the aforementioned statistical variability. We can, in fact, interpret the network as a circuit of non-linear resistances that can conceptually be reduced to a single resistor, whose *I–V* characteristic depends on the node positions. Thus, prescribing the same current gives origin to different potential drops for different node placements. The sensitivity of the threshold condition to the particular location of the nodes is significantly reduced for longer cutoff distances, becoming of the order of ∼1% for V_{th} and ∼13% for I_{th} in the limit of large r_{cut} (see Table II). Experimentally, subsequent amorphizations of the same device after burn-in[16] or, equivalently, the amorphization of a number of geometrically identical devices, may originate different configurations of defects inside the amorphous matrix, producing a dispersion in V_{th} and I_{th}.

In conclusion, this letter presents a 3D network model accounting for the effect of disorder in amorphous phase-change materials. A set of balance equations coupled to the Poisson equation is employed to predict the *I–V* relationships; good agreement with recent experimental data is found for phase-change memory devices probing just hundreds of cubic nanometers of GST.[16]

The microscopic analysis of inter-node transitions correlates the threshold-switching event with the concentration of current into preferential hot-electron paths, shunting the electrodes. The stochastic network generation allows one to study statistical fluctuations in the threshold, also pointing out that this variability depends on the interplay between the network configuration and the typical interaction length of the nodes. The model can provide a valid, fast simulation and design tool for more complex geometries, as the node distribution criteria can be adapted to represent arbitrary device architectures.

Finally, we remark that the high-energy conductive paths may produce crystalline nucleation sites. This suggests a thermal extension of the model, possibly establishing a link between the points of view of the crystalline filament and the purely electronic switching theories.

The Italian authors acknowledge valuable discussions with Dr. F. Buscemi and Professor L. Zanni and financial support from the Intel Corporation under Contract No. 3477131/2011. The Illinois authors acknowledge support from the U.S. Office of Naval Research (Grant No. N00014-10-1-0853) and the U.S. National Science Foundation (Grant No. ECCS 1002026).

[1]S. R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968).
[2]M. Wuttig and N. Yamada, Nature Mater. **6**, 824 (2007).
[3]J. Yao, Z. Sun, L. Zhong, D. Natelson, and J. M. Tour, Nano Lett. **10**, 4105 (2010).
[4]H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. Chen, and M.-J. Tsai, Proc. IEEE **100**, 1951 (2012).
[5]A. Shindome, Y. Doioka, N. Beppu, S. Oda, and K. Uchida, Jpn. J. Appl. Phys. **52**, 04CN05 (2013).
[6]A. Sebastian, A. Pauza, C. Rossel, R. M. Shelby, A. F. Rodríguez, H. Pozidis, and E. Eleftheriou, New J. Phys. **13**, 013020 (2011).
[7]R. G. Neale and J. Aseltine, IEEE Trans. Electron Devices **20**, 195 (1973).
[8]D. Kau, S. Tang, I. Karpov, R. Dodge, B. Klehn, J. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee, T. Langtry, K. wei Chang, C. Papagianni, J. Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro, and G. Spadini, IEDM - Tech. Dig. **2009**, 1–4.
[9]H. -S. P. Wong, S. Raoux, S. Kim, J. Liang, J. P. Reifenberg, B. Rajendran, M. Asheghi, and K. E. Goodson, Proc. IEEE **98**, 2201 (2010).
[10]K. Tsendin, Phys. Status Solidi B **246**, 1831 (2009).
[11]D. Ielmini, Phys. Rev. B **78**, 035308 (2008).
[12]E. Piccinini, A. Cappelli, F. Buscemi, R. Brunetti, D. Ielmini, M. Rudan, and C. Jacoboni, J. Appl. Phys. **112**, 083722 (2012).
[13]C. Jacoboni, E. Piccinini, F. Buscemi, and A. Cappelli, Solid-State Electron. **84**, 90 (2013).
[14]M. Nardone, V. G. Karpov, D. C. S. Jackson, and I. V. Karpov, Appl. Phys. Lett. **94**, 103509 (2009).
[15]F. Xiong, A. D. Liao, D. Estrada, and E. Pop, Science **332**, 568 (2011).
[16]F. Xiong, M.-H. Bae, Y. Dai, A. D. Liao, A. Behnam, E. A. Carrion, S. Hong, D. Ielmini, and E. Pop, Nano Lett. **13**, 464 (2013).
[17]N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon, Oxford, 1979).
[18]D. Ielmini and Y. Zhang, Appl. Phys. Lett. **90**, 192102 (2007).
[19]D. Ielmini and Y. Zhang, J. Appl. Phys. **102**, 054517 (2007).
[20]G. B. Beneventi, L. Guarino, M. Ferro, and P. Fantini, J. Appl. Phys. **113**, 044506 (2013).
[21]See supplementary material at http://dx.doi.org/10.1063/1.4819097 for the complete derivation and discussion of the equations and the solution scheme, the estimation of the amorphous bit size, the comparison of the results from the present model with those obtained with the HD model, and three animations describing the microscopic evolution of the network as a function of the prescribed current.
[22]P. Ramachandran and G. Varoquaux, Comput. Sci. Eng. **13**, 40 (2011).
[23]D. Adler, M. S. Shur, M. Silver, and S. R. Ovshinsky, J. Appl. Phys. **51**, 3289 (1980).

Supplementary materials for Conductive preferential paths of hot carriers in amorphous phase-change materials

Andrea Cappelli,[1] Enrico Piccinini,[2, a)] Feng Xiong,[3] Ashkan Behnam,[3] Rossella Brunetti,[1] Massimo Rudan,[2] Eric Pop,[3] and Carlo Jacoboni[1]

[1)] *Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Via Campi 213/A, I-41125 Modena, Italy*

[2)] *"E. De Castro" Advanced Research Center on Electronic Systems (ARCES), Università di Bologna, Via Toffano 2/2, I-40125 Bologna, Italy*

[3)] *Department of Electrical and Computer Engineering, Micro and Nanotechnology Lab University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

[a)] Author to whom correspondence should be addressed. Electronic mail: enrico.piccinini@unimore.it

I. DERIVATION OF THE CHARGE-FLUX AND OF THE ENERGY-FLUX BALANCE EQUATIONS

Let us consider a network of $N+2$ sites as in Figure S1. The two nodes labeled 0 and $N+1$ are special nodes playing the role of contacts for the device.

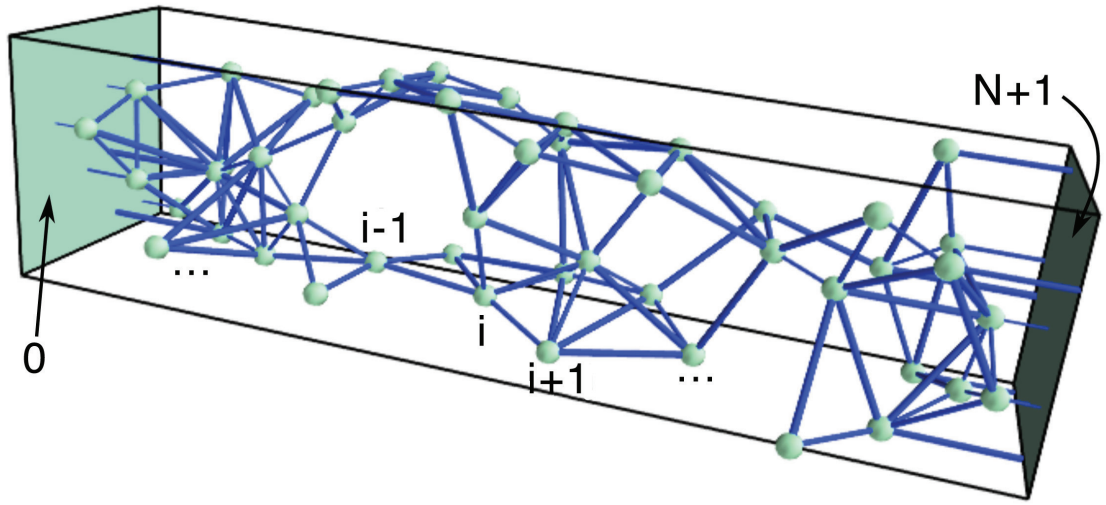

FIG. S1. Schematics of the simulated domain. The shaded surfaces represent the source and the drain contact, respectively. Distances between contacts and internal nodes are calculated along the perpendicular line.

We want to write balance equations for the charge- and energy- fluxes among these nodes. The symbols have the same meaning as in the main text. Let $\mathrm{d}t$ be an infinitesimal time during which a fraction of the population of each node transfers to other nodes. We define the following quantities with respect to the generic i-th node:

1. population flowing out of node: $\mathrm{d}t\left[n_i\sum_j S_{ij}+\frac{I}{(-q)}\delta_{N+1,i}\right]$;
2. population flowing into node: $\mathrm{d}t\left[\sum_j n_j S_{ji}+\frac{I}{(-q)}\delta_{0,i}\right]$;
3. energy flowing out of node: $\mathrm{d}t\,\varepsilon_i\left[n_i\sum_j S_{ij}+\frac{I}{(-q)}\delta_{N+1,i}\right]$;
4. energy flowing into node: $\mathrm{d}t\Big\{\sum_j n_i S_{ji}\Big[\varepsilon_j-q(\varphi_j-\varphi_i)\Big]+\frac{I}{(-q)}\varepsilon_i\delta_{0,i}\Big\}$;
5. energy to relax: $\varepsilon_i^R=\mathrm{d}t\sum_j n_j S_{ji}\Big[(\varepsilon_j-\varepsilon_{i,eq})-q(\varphi_j-\varphi_i)\Big]+\Big[n_i\left(1-\mathrm{d}t\sum_j S_{ij}\right)-\mathrm{d}t\frac{I}{(-q)}\delta_{N+1,i}\Big](\varepsilon_i-\varepsilon_{i,eq})$;

6. relaxed energy: $\varepsilon_i^R f_R = \varepsilon_i^R \Big[1 - \exp(-\mathrm{d}t/\tau_R)\Big]$.

where the summations do not take the case $i = j$ into account. By means of these definitions we can write the population balance for the i-th node:

$$n_i(t + \mathrm{d}t) = n_i(t) - \mathrm{d}t \left[n_i \sum_j S_{ij} + \frac{I}{(-q)} \delta_{N+1,i} \right] + \mathrm{d}t \left[\sum_j n_j S_{ji} + \frac{I}{(-q)} \delta_{0,i} \right], \tag{S1}$$

and the energy balance:

$$\begin{aligned} n_i(t + \mathrm{d}t)\varepsilon_i(t + \mathrm{d}t) = n_i(t)\varepsilon_i(t) - \mathrm{d}t\, n_i\varepsilon_i \sum_j S_{ij} + \mathrm{d}t \sum_j n_j S_{ji} \Big[\varepsilon_j - q(\varphi_j - \varphi_i)\Big] + \\ + \mathrm{d}t \frac{I}{(-q)} \varepsilon_i \Big[\delta_{0,i} - \delta_{N+1,i}\Big] - \Bigg\{ \mathrm{d}t \sum_j n_j S_{ji} \left[(\varepsilon_j - \varepsilon_{eq}) + q_e(\varphi_i - \varphi_j)\right] + \\ + \left[n_i \Big(1 - \mathrm{d}t \sum_j S_{ij}\Big) - \mathrm{d}t \frac{I}{(-q)} \delta_{N+1,i} \right] (\varepsilon_i - \varepsilon_{i,eq}) \Bigg\} \times \Big[1 - \exp(-dt/\tau_R)\Big]. \end{aligned} \tag{S2}$$

The l.h.s. and the last factor at the r.h.s. of Eq. S2 can be expanded to the first order to yield:

$$\begin{aligned} n_i(t + \mathrm{d}t)\varepsilon_i(t + \mathrm{d}t) &= \left[n_i(t) + \frac{\mathrm{d}n_i}{\mathrm{d}t}\mathrm{d}t \right] \left[\varepsilon_i(t) + \frac{\mathrm{d}\varepsilon_i}{\mathrm{d}t}\mathrm{d}t \right] \\ &= n_i(t)\varepsilon_i(t) + n_i(t)\frac{\mathrm{d}\varepsilon_i}{\mathrm{d}t}\mathrm{d}t + \frac{\mathrm{d}n_i}{\mathrm{d}t}\varepsilon_i(t)\mathrm{d}t + \frac{\mathrm{d}n_i}{\mathrm{d}t}\frac{\mathrm{d}\varepsilon_i}{\mathrm{d}t}\mathrm{d}t^2, \end{aligned} \tag{S3}$$

and

$$f_R = 1 - \exp(-\mathrm{d}t/\tau_R) \approx 1 - 1 + \frac{\mathrm{d}t}{\tau_R} = \frac{\mathrm{d}t}{\tau_R}. \tag{S4}$$

Neglecting the infinitesimal terms, and remembering that $\mathrm{d}\left(n_i\varepsilon_i\right)/\mathrm{d}t = n_i\mathrm{d}\varepsilon_i/\mathrm{d}t + \varepsilon_i\mathrm{d}n_i/\mathrm{d}t$, Eqs. S1 and S2 can be rewritten as:

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = -n_i \sum_j S_{ij} - \frac{I}{(-q)}\delta_{N+1,i} + \sum_j n_j S_{ji} + \frac{I}{(-q)}\delta_{0,i} = 0, \tag{S5}$$

$$\begin{aligned} \frac{\mathrm{d}(n_i\varepsilon_i)}{\mathrm{d}t} = &- n_i\varepsilon_i \sum_j S_{ij} + \sum_j n_j S_{ji} \Big[\varepsilon_j - q(\varphi_j - \varphi_i)\Big] + \frac{I}{(-q)}\varepsilon_i \Big[\delta_{0,i} - \delta_{N+1,i}\Big] + \\ &- n_i \frac{\varepsilon_i - \varepsilon_i^*}{\tau_R} = 0, \end{aligned} \tag{S6}$$

where we have imposed $\mathrm{d}n_i/\mathrm{d}t = 0$ and $\mathrm{d}(n_i\varepsilon_i)/\mathrm{d}t = 0$ since we are interested in the steady-state solution. With little, straightforward manipulation Eqs. S5 and S6 become Eqs. 2 and

3 of the main text. The latter expression can then be recast into a more compact and handy formula by multiplying Eq. S5 times ε_i and substituting it into Eq. S6:

$$\varepsilon_i = \varepsilon_{i,eq} + \tau_R \sum_j \frac{n_j}{n_i} S_{ji} \Big[\varepsilon_j - \varepsilon_i - q(\varphi_j - \varphi_i) \Big] \tag{S7}$$

II. SOLUTION SCHEME

Formulas S5 and S7 represent two sets of equations whose unknowns are the nodal populations n_i's, electrostatic potentials φ_i's, and energies ε_i's. Since the number of nodes is $N+2$, we need $3(N+2)$ independent equations to solve the problem.

Equation S5 represents the well-known Kirchoff's Current Law applied to each node. Since any network, despite its complexity, can always be reduced to a single resistor, only a set of $N+1$ useful equations can be derived, the remaining one not adding any further information.

We already pointed out that nodes 0 and $N+1$ play the role of the two contacts. They are assumed to be planar, in the sense that distances between one of them and any of the internal nodes is calculated as the minimum distance between the internal node and the appropriate face of the device. Additional boundary conditions apply to the contacts. In particular, we can assume that their population is always the equilibrium population ($n_0 = n_{N+1} = n_{eq}$). Moreover, we can also suppose that carriers enter the device from the source contact being *cold*, i.e., $\varepsilon_0 = \varepsilon_{eq}$. On the other side of the device the drain contact is supposed to receive *hot* carriers, and its nodal energy depends on the efficiency of the relaxation mechanism at the contact interface. For a maximally efficient energy transfer we would suppose $\varepsilon_{N+1} = \varepsilon_{eq}$; otherwise, this value is unknown and must be determined by solving the problem. For the sake of truth, since the fraction of carriers entering back into the device from the drain contact is negligible at any conduction regime, very similar results are found in both cases. In order to discuss the problem under the most general conditions we consider in the followings the energy of the drain contact as unknown.

The considerations above reduce the number of unknowns to $3N+3$, a number that can be further reduced to $3N+2$ as we recall the arbitrariness of the electrostatic potential, which allows to set, for instance, $\varphi_0 = 0$. From Eqs. S5 and S7 only $2(N+1)$ equations are

available. The missing N relationships stem from the Poisson equation:

$$\nabla^2\varphi(x,y,z) = -\frac{\rho(x,y,z)}{\epsilon}, \tag{S8}$$

which links the electrostatic potential to the charge, thus the population, of each node.

A finite-element implementation of the latter over a grid of M points ($M \gg N$) is adopted, and a buffer region surrounds laterally the device to avoid the influence of artificial boundary conditions. If the electrostatic potentials of some nodes are known from an independent source of information, the finite-element solution of Eq. S8 yields the charge on these nodes and the electrostatic potential elsewhere, providing the required N supplementary relationships.

Algorithm 1 Finds the voltage drop, for a given input current I

Require: Device and material parameters, network topology, prescribed input current I

while $\left\| n_i^{(s+1)} - n_i^{(s)} \right\| > \Delta_n$ **do**
  $s \leftarrow s+1$
  while $\left\| f(\varphi_i, \varepsilon_i, n_i^{(s)}) \right\| > \Delta_{tol}$ **do**
    NRP for Eq. S5 at fixed $(\varepsilon_i, n_i^{(s)}) \rightarrow \varphi_i$
    NRP for Eq. S7 at fixed $(\varphi_i, n_i^{(s)}) \rightarrow \varepsilon_i$
  end while
  PS using $(\varepsilon_i, \varphi_i) \rightarrow n_i^{(s+1)}$
end while
return $V = \varphi_{N+1}^{(s+1)} - \varphi_0^{(s+1)}$

Equations S5 and S7 are strongly non-linear due to the exponential form of S_{ij} and S_{ji} (see Eq. 1 in the main text), and also depend on the populations, that come from the Poisson equation, which needs the electrostatic potentials of the nodes as input values. Iterative methods of a different kind must be applied for finding potentials and energies on one side, and for populations on the other side. The Newton-Raphson procedure (NRP) for sets of non-linear equations has been implemented in the first case, assuming a constant population. It proved to be convenient to address the two sets S5 and S7 separately, keeping in the first case the energies of the nodes fixed to obtain an updated solution for φ_i, and keeping the new potentials fixed in the second case in order to update ε_i. These steps are cycled until a global convergence is achieved for both potentials and energies (overall NRP,

or oNRP). Once the oNRP phase is finished, the population of each node is updated by evaluating the charge Q_i of the nodes through a Poisson step (PS), and by applying the relationship $n_i = Q_i/(-q) + n_{eq}$. Pairs of oNRP's and PS's are repeated until all of the unknowns are determined within their desired precision (see Algorithm 1). The voltage drop $V = \varphi_{N+1} - \varphi_0$ is the final outcome of the method for any prescribed input current I.

III. ESTIMATE OF THE DIMENSIONS OF GST-CNT DEVICES

The model has been validated by checking calculated current-voltage characteristics against experimental data of carbon-nanotube-contacted self-aligned $Ge - Sb - Te$ memory devices (CNT-GST), fabricated as in Ref. 1. In order to achieve reliable comparisons, it is important to feed the model the appropriate size of the chalcogenide volume.

The manufacturing process of CNT-GST devices implies an inherent random distribution of their lengths[1,2]. A first-order estimate of these values stems from the observation that the onset of a negative differential resistance is obtained when a threshold field F_{th} is reached. Literature values for such a field report typically $F_{th} \approx 0.5$—1 MV/cm (50—100 V/μm) for CNT-GST devices[1]. For the present case, we have assumed the intermediate field $F_{th} \sim 0.9$ MV/cm (90 V/μm) for all cases considered. The simple calculation V_{th}/F_{th} leads to the three estimated lengths 40 nm, 81 nm and 91 nm for the three devices included in the analysis.

An estimate of the cross-sectional area is more complicated, as the active chalcogenide spot is expected to extend on a section larger than the CNT area[1]. Typical values for CNT-GST devices lie in the range 90—180 nm^2. In the calculations we have assumed a constant 10×10 nm^2 cross section for all devices, also considering the trade-off with the computational burden. This value may sometime underestimate the real effective cross-section, as in the case of devices featuring the highest threshold currents. As we have also commented in the main text, such underestimation may be compensated by a larger variability of the relaxation time τ_R. However, different choices of the threshold field F_{th} or of the cross-sectional area would affect results only quantitatively without influencing the physics underneath.

IV. COMPARISON WITH THE MICROSCOPIC HYDRODYNAMC-LIKE MODEL

The network model and the microscopic HydroDynamic (HD) model[3] share the same physical background for charge transport. It is therefore possible to qualitatively compare the parameters used in the two models to fit the I–V characteristics. It is important to recall that the HD model is a continuum-medium model and the momenta of transport equations are obtained integrating a distribution function, whereas the network model makes use of a spatially-discrete approach and considers directly the average energy of the carriers. As a consequence, the common background of the models makes the comparison of the parameters possible only from a qualitative standpoint.

The experimental I–V characteristic of device L40 (see main text) has been fitted also with the HD model (Fig. S2); parameters are listed in Table TI along with those obtained with the same model for traditional mushroom cells[3]. For comparison purposes, we have also reported the parameters of the statistical best-fit of device L40 with the network model.

The HD model considers a density of states of traps Γ populated according to the distribution function. By integrating it over the energy and the device volume we obtain a carrier concentration of the order of $1—5 \times 10^{19}$ cm^{-3}, when $\Gamma \sim 10^{20}$ cm^{-3} eV^{-1}. In the network model, a node can in principle host an arbitrary number of carriers. If we assume the simplifying hypothesis of one carrier per node under the initial equilibrium condition, we obtain the carrier concentration $\sim 10^{19}$ cm^{-3}, which well compares to that obtained with the HD model for device L40.

TABLE TI. Parameters of the HD mode for device L40 and for a mushroom cell with cross-sectional area 1000 nm^2 and length 27 nm, and for the network model for device L40 (from the main text).

HD model			Network model	
	L40	Ref. 3		L40
Γ (cm^{-3} eV^{-1})	0.4×10^{20}	1.18×10^{20}	n (cm^{-3})	1.2×10^{19}
τ_0 (fs)	50	12	τ_0 (fs)	49.5
τ_R (fs)	17	78	τ_R (fs)	56.5
Δz (nm)	3.2	7	r_{cut} (nm)	6
			ℓ (nm)	2.9

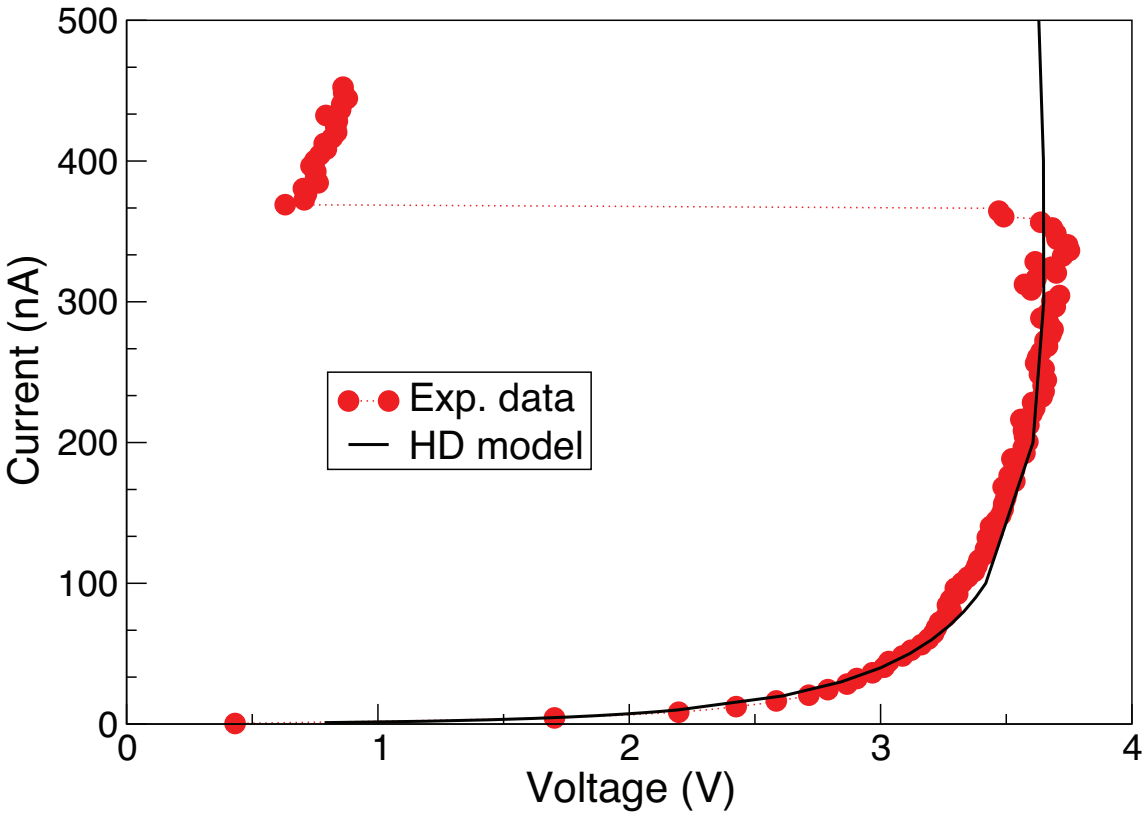

FIG. S2. I–V characteristic (line) obtained with the HD model and the parameters listed in Table TSTI and experimental data for device L40 (dots).

Next, we find similar values also for the characteristic time of the transition process τ_0 and the energy-relaxation time τ_R, that are estimated by both models in the range 10-100 fs. This result confirms once again the common nature of the physical processes of the transport framework, with numerical differences depending only on the specific details of their implementation.

A slightly more complicated reasoning stands behind parameter Δz in the HD model, which makes the comparison with distances r_{cut} and ℓ in the network model less direct. Following Ref. 4, Δz is the average transition distance and is basically determined fitting the slope of the exponential portion of the I–V curve when one assumes that the top of the energy barrier governing the transition is at the midpoint of the transition line. In the network model, these two ideas have been split and the two parameters r_{cut} and ℓ must be determined independently. Going into further details, r_{cut} represents the maximum

transition distance. Given the constraint of a minimum distance r_{min} between two nodes, which is a condition to avoid node overlap, and assuming a uniform transition probability within a spherical shell from r_{min} to r_{cut}, the average transition distance can be roughly estimated as $3/4 \times (r_{cut}^4 - r_{min}^4)/(r_{cut}^3 - r_{min}^3)$ which is up to 7% larger than $3/4 r_{cut}$ for $r_{min} \lesssim r_{cut}/2$. For the test case of device L40, according to the reasoning above, we can estimate a transition distance of 4.7 nm for $r_{min} \sim r_{cut}/2$, a value 50% larger than what predicted by the HD model.

As for the parameter ℓ (effective correction to the height of the energy barrier), we recall that in the network model this correction depends on the distance of the nodes, being ℓ its maximum value. In other words, for transitions where $r_{ij} < 2\ell$ we assume that the maximum of the energy barrier is located at the midpoint of the transition line, i.e., $r_{ij}/2$, as in the HD model; on the contrary, for longer transitions we assume a fixed position at distance ℓ. Though the procedure is tricky, we can still indirectly estimate ℓ from to the slope of the exponential portion of the I–V characteristic by weighting short and long transitions. The ratio ℓ/r_{cut} close to 0.5 indicates that the hypothesis on the position of the maximum of the energy barrier in the HD model is confirmed also in the network model.

V. MOVIES

Three movies are provided along with this work. Each of them shows the microscopic network for a particular different random node configuration, with color codes representing the energies ε_i on the nodes and the intensities I_{ij} of the net currents among the nodes, normalized to the total current I. Next to the movie, an I–V characteristic curve is reported, with the moving dot showing, for each configuration frame, its corresponding values of voltage V and current I. The 3D visualization software Mayavi[5] was employed to produce these movies and the screenshots in the main text.

Movie #1 is the movie from which the screenshots of the main text are extracted. For low currents, no unique continuous path is observed to directly connect the electrodes.This movie clearly highlights the effect of "channeling" when higher currents are prescribed: even for a current level of a few nanoamperes, the path in the upper part of the device begins to get enriched. For slightly higher I, about $\sim$ 1—50 nA, it becomes evident that such current increase in the main path happens at the expense of a flow decrease through other

secondary links, particularly the two consecutive ones in the bottom part of the sample.In the last frames, the "channeling" is almost total, and a clear path connects the electrodes while little or no flow is observed in the other lines.

Movie #2 is obtained for a value of $r_{cut} = 8$ nm, higher than the 6 nm used for the first movie. This has a clear effect: the current can be spread among more available connections, thus a more uniform low-current picture is obtained; initially no paths are strongly favored over the others simply on a geometrical basis. Anyway, a region of low node density in the middle of the device (y = 15—22 nm) makes it necessary to have slightly above-average values of transition rates, in order to sustain the prescribed current, since not many parallel viable paths are available in that portion of the sample. For higher currents ($I \geq 10$ nA) the current distribution starts to change rapidly. Electrons in the middle region heat up as a result of the higher "current densities" associated to it, thus allowing even more current to flow, and triggering a positive feedback. Indeed, this area is the point where the formation of a single preferential path begins for this device. In the subsequent frames, the path spreads from the center toward both contacts, linking high-energy nodes, and finally connects the device end-to-end for $I \sim 7$ nA.

Movie #3 shows the case of a 91-nm-long device, with $r_{cut} = 6.5$ nm. When r_{cut} is small compared to the device length, segmentation of the device is present, i.e., regions densely populated with nodes intertwined with others with a very limited number of nodes. For the case at hand, three highly-populated segments can be identified, two close to the two contacts and one in the middle of the device. As the current is increased from low values to the threshold condition, one path in the poorly populated regions progressively drains almost the entire current. This situation makes transport within highly-populated segments independent from each other, as if long devices were consecutive sequences of short ones. Moreover, it is also worth noticing that the preferred route leading to threshold switching can be further modified by the transport conditions. In particular, we point out that before the threshold condition there are two competitive paths converging to a node at $y \sim 45.5$ nm that gathers the current from the surrounding paths. After switching, the interplay between them is such that only one path continues to increase its current, draining it from one link that contributed to create the preferential path up to that moment. This is likely due to the strongly non-linear nature of the balance equations.

REFERENCES

[1]F. Xiong, M.-H. Bae, Y. Dai, A. D. Liao, A. Behnam, E. A. Carrion, S. Hong, D. Ielmini, and E. Pop, Nano Lett. **13**, 464 (2013).

[2]F. Xiong, A. D. Liao, D. Estrada, and E. Pop, Science **332**, 568 (2011).

[3]E. Piccinini, A. Cappelli, F. Buscemi, R. Brunetti, D. Ielmini, M. Rudan, and C. Jacoboni, J. Appl. Phys. **112**, 083722 (2012).

[4]D. Ielmini and Y. Zhang, J. Appl. Phys. **102**, 054517 (2007).

[5]P. Ramachandran and G. Varoquaux, Comp. Sci. Eng. **13**, 40 (2011).

5.1 Effect of disorder degree

The information provided by the model about the detailed situation of the network gives a chance to qualitatively analyze how the network topology can affect the electrical properties of a device. In particular, it is possible to qualitatively assess the effect of different degrees of "network disorder".
In this context, the concept of disorder is thought of as the absence of a regular and somehow periodic spatial distribution of the network nodes while, on the contrary, such regularity is present in an ordered network. In particular, the ordered case is naturally associated to the presence of several parallel paths, electrically equivalent to each other, through which the current can evenly spread while flowing into the device. Gradually larger deviations from this ideal condition can be introduced in the system to study their consequences on charge transport.

A hexagonally-symmetric prismatic configuration was chosen as the initial ordered case (with ∼5 nm interplane node distances and ∼4 nm intraplane node distances). Starting from such condition, displacement vectors were randomly generated for each node, according to a 3D isotropic Gaussian probability density. The standard deviation σ_d of such distribution was used to determine the typical magnitude of the displacement, and thus the level of deviation from the ordered case.
Figure 5.1 depicts three sets of nodes (shown from both side and front views), the first one being the beginning ordered one and the subsequent two being the perturbed cases, with $\sigma_d = 1.0$ nm and $\sigma_d = 4.0$ nm, respectively. With higher disorder, the initial distribution of nodes becomes increasingly difficult to recognize, and the resulting parallel equivalent paths are gradually lost.
Figure 5.2 shows the IV curves corresponding to several configurations obtained for σ_d going from 0.0 nm to 4.0 nm, including the three cases from Figure 5.1. Two qualitative trends are observed: at first, for σ_d up to 1.0 nm, a slight conductivity increase is found. This is somewhat counter-intuitive, since the idea of disorder is commonly associated with lower conductivity. We notice that, due to FEM grid constraints, the in-plane hexagons are not perfectly regular. As a consequence, the initial configurations might not be the optimal one for conduction, and small displacements of the nodes could have made the system closer to such situation. However, for higher values of σ_d, the conductivity significantly decreases and the threshold voltage of the device steadily increases, both behaviors being more intuitive and physically sound for a less regular system. As described earlier ([27], reported above), the relation between the transport properties and the network topology is mediated by the cutoff distance parameter r_{cut}. Its influence can

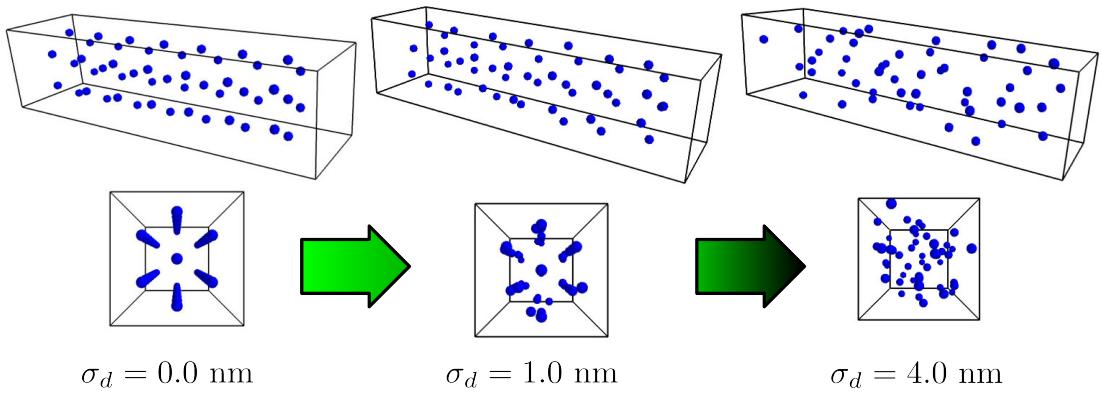

Figure 5.1: Visualization of the increasing network disorder for $\sigma_d = 0.0$,1.0,4.0 nm. For each disorder level, the upper picture is a side view of the network nodes, the lower one a front view.

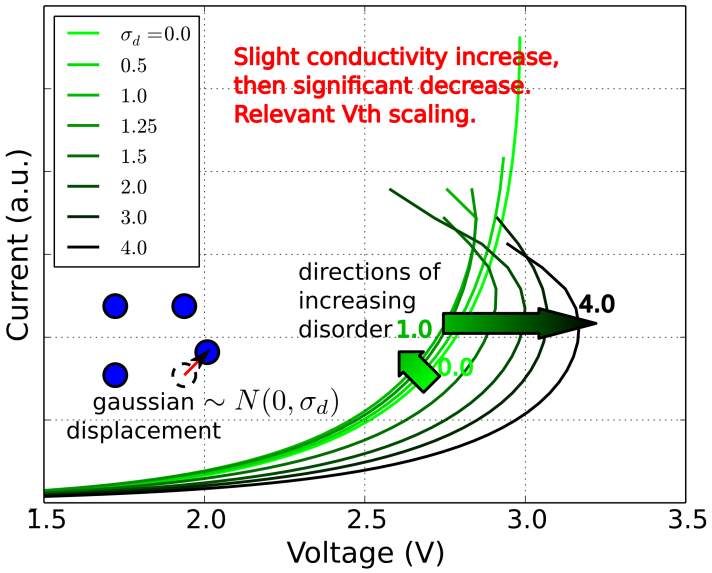

Figure 5.2: Effect of the amount of disorder on conduction properties. The curves are related to network configurations generated perturbing an initial ordered network with random Gaussian displacements (see the inset scheme), with an increasing standard deviation σ_d. Darker-green curves come from more disordered situations (i.e. higher σ_d). The lightest-green curve is for the initial ordered configuration. The two arrows (same color code) indicate the directions of increasing disorder, the shortest one for $\sigma_d = 0.0 \rightarrow 1.0$ nm, the other for $\sigma_d = 1.0 \rightarrow 4.0$ nm.

be complicate to assess, since it involves the analysis of the particular distances between pairs of network nodes, which can be brought into/out of the r_{cut} radius by a particular random displacement. This can actually lead to counter-intuitive behavior such as the initial disorder-induced conductivity increase.

5.2 Thermal analysis and local crystallization

In the following we present a thermal extension of the model described above [27]. This is somehow similar to the one adopted in [103, 102], although in that case a continuum approach is employed.
The thermal analysis is fulfilled by coupling the charge-balance, energy-balance and Poisson equations to the Fourier heat transport equation. This allows to both assess the conductivity-enhancing effect of the increasing lattice temperature and study the onset of the crystallization event. A critical crystallization temperature $T_c \sim 420$ K is adopted for the phase change, and a link between two network nodes is considered crystallized when the average temperature along its length is greater than T_c. Such event is modeled by a change of the set of values used to parametrize the node-to-node transition rates, in such a way that conductivity increases by about two orders of magnitude.
The effect of the interplay among material properties such as the thermal conductivity κ and the energy relaxation characteristic time τ_R is assessed in relation with the occurrence of the Ovonic and crystallization switching events. The simulation results are then compared with IV curves from CNT-GST devices, and device-to-device variability is analyzed for different boundary conditions, corresponding to two different device architectures (lateral and mushroom-like).
The topic is presented through the inclusion of the original paper [92].

Novel 3D random-network model for threshold switching of phase-change memories

Enrico Piccinini[1], Andrea Cappelli[2], Feng Xiong[3,4], Ashkan Behnam[3], Fabrizio Buscemi[1], Rossella Brunetti[2], Massimo Rudan[1], Eric Pop[3,4] and Carlo Jacoboni[2]

[1]ARCES Research Center and DEI, Università di Bologna, Bologna, Italy
[2]Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Italy
[3]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, IL, USA
[4]now with Department of Electrical Engineering, Stanford University, CA, USA
phone +39 059 205 5292, fax +39 059 205 5616, email: enrico.piccinini@unimore.it

Abstract

The onset of crystallization in phase-change memory devices is studied by simulating an initially amorphous sample through a disordered network of localized states. The transport of charge and electron energy is self-consistently coupled to the Poisson and the Fourier heat equations, so that crystallization sites are found at the nanoscale. Results show how Ovonic switching and crystallization are both correlated to the formation of hot-carrier conduction paths, and the conditions for the occurrence of these phenomena are investigated. The model is then validated against data from ultra-scaled carbon-nanotube-contacted devices. Device-to-device variability of macroscopically identical devices is also analyzed.

Introduction

The electronic industry is nowadays designing for new non-volatile memory solutions beyond the FLASH era, showing better performances and a reduced energy consumption with respect to the state-of-the-art technology. Phase-change memories (PCM) represent the top candidate for the next generation of non-volatile memories and have recently been released to the market (1). Their working principle relies on encoding a bit of information through the reversible resistance change of a nanometric volume of a chalcogenide compound, induced by the amorphization or the crystallization of the chalcogenide material itself.

PCM feature good operation speed and endurance (2), they are suitable for cross-point three-dimensional (3D) stacking (3), and have an extremely promising scaling outlook (4). These qualities can pave the way for the creation of very low-power devices whose energy consumption for set and reset operations lies in the femto-Joule range.

Further scaling of current PCM devices, as well as the design of ultra-scaled non-conventional prototypes, undoubtedly require a solid knowledge of the underlying physics and the availability of fast and reliable simulation tools for the amorphous phase. In particular, few-carrier phenomena and local non-uniformities are critical challenging issues to obtain reliable devices featuring either Ovonic- or Memory-Threshold Switching (OTS or MTS, respectively), or both.

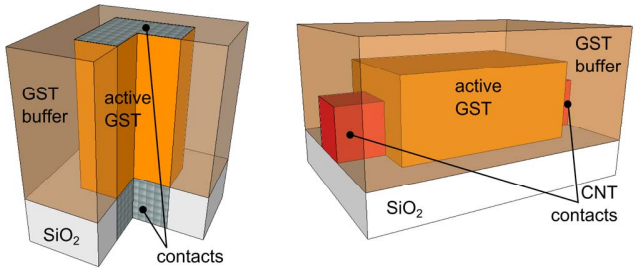

Figure 1 – Schematics of the simulated devices: a traditional "mushroom-like" cell (left) and the novel device of (4) using carbon nanotubes as contacting leads (right). Lateral adiabatic boundary conditions apply in both cases; the SiO_2 bottom layer is kept at 300 K.

The transport model presented in this work addresses these issues and provides a physical insight on the conduction process at the nanoscale level. More specifically, the transport picture includes local thermal and electrical effects and the random nature of amorphous materials in a full 3D computational domain. The outcome of the simulations demonstrates how OTS and MTS are both connected to the onset of preferential paths of hot carriers that also play the role of crystallization sites. The model is also predictive about the variability of the threshold conditions of batches of geometrically identical devices. Simulations are compared to experimental results for memory prototypes using carbon nanotubes as contacting leads.

The model

Even though some aspects are still a debated question (5), conduction in amorphous chalcogenides has been ascribed to trap-limited transport (6). The latter is conveniently described as sequences of detrapping, propagation and trapping events, which are modeled via the Poole and/or the Poole-Frenkel formulas. Following such approaches, continuum-medium analytical models were proposed, either in simple one-dimensional (1D) or in more realistic 3D cases (7-9).

Very recently, Cappelli *et al.* have proposed a network approach that goes beyond both the 1D and the continuum-medium approximations, and allows one to investigate the conduction at a microscopic scale (10). Such a model consti-

tutes the background of the present work, and we address the reader to (10) for its detailed description. Schematics of the simulated devices used in the next sections are reported in Fig. 1.

A number N of scattering centers is placed at random in the active part of the simulation domain. Each center hosts a population n_i, whose average energy is e_i, and it also features an electrostatic potential φ_i and a local temperature T_i. These centers exchange carriers with each other according to a modified Poole transition rate S_{ij}:

$$S_{ij} = \frac{1}{\tau_0} f(r_{ij}) \exp\left(-\frac{E_c - e_i}{k_B T_i}\right) \exp\left[-\frac{q(\varphi_i - \varphi_j)\ell}{r_{ij} k_B T_i}\right], \quad (1)$$

with $f(r_{ij})$ a function of the inter-center distance r_{ij} that, for simplicity, is assumed 1 when r_{ij} is within the interaction distance r_{cut} and 0 otherwise. In (1), τ_0 is a characteristic transfer time, E_c is the bottom of the conduction band, k_B is the Boltzmann constant, q is the electron charge and ℓ is an average effective distance that takes into account the effect of the electric field on the height of the barrier separating the scattering centers i and j. By means of (1), we write the charge and the energy flux balances at any scattering center j in the steady state condition, after taking into account the energy gain due to the local electric field and the losses associated to phonon interactions:

$$\frac{I}{(-q)}\delta_{0,j} + \sum_{i\neq j} n_i S_{ij} = \sum_{i\neq j} n_j S_{ji} + \frac{I}{(-q)}\delta_{N+1,j}, \quad (2)$$

$$\sum_{i\neq j} n_i S_{ij}\left[e_i - q(\varphi_i - \varphi_j)\right] + \frac{I}{(-q)} e_j \delta_{0,j} = e_j\left[\sum_{i\neq j} n_j S_{ji} + \frac{I}{(-q)}\delta_{N+1,j}\right] + n_j \frac{e_j - e_{j,eq}}{\tau_R}, \quad (3)$$

where τ_R is the energy relaxation time constant. The Kronecker's δ-functions prescribe the current I at the two contacts, which are labeled as 0 and $N+1$. The eq index indicates quantities evaluated under equilibrium conditions.

Equations (2) and (3) are coupled to the Poisson equation:

$$\nabla \cdot (\varepsilon \nabla \varphi) + Q_j = 0, \quad (4)$$

where $Q_j = (-q)(n_j - n_{j,eq})$ at the scattering centers j and 0 otherwise, and to the Fourier heat equation:

$$\nabla \cdot (\kappa \nabla T) + W_{ij} = 0, \quad (5)$$

with $W_{ij} = q(n_i S_{ij} - n_j S_{ji})\left[e_i - e_j - q(\varphi_i - \varphi_j)\right]$ uniformly distributed along the line connecting centers i and j and 0 otherwise. Parameters ε and κ are the dielectric constant and the thermal conductivity of the materials, respectively. Equations (4) and (5) are solved using the finite-element method; notice that the definitions of Q_j and W_{ij} imply that the scattering centers are a subset of the nodes of the finite element grid.

The temperature distribution obtained from the solution of eq. (5) determines whether any line connecting two scattering centers or any spot of the simulated domain crystallizes. The crystallization temperature T_c = 420 K is assumed for PCM cells featuring GST-225 as the active chalcogenide material, meaning that any grid block or connection between scattering centers with an average temperature greater than T_c is assumed crystalline in the subsequent steps of the simulation. Crystalline connections are simulated by changing the parameters of eq. (1) in order to reduce their resistances of two decades; crystalline portions of the domain also have a larger thermal conductivity. However, the position of the very first crystallization events may induce in the other connections temporary extra currents, which are not reproduced by the present steady-state model. As a consequence, simulations are terminated as soon as the first crystallization events occur.

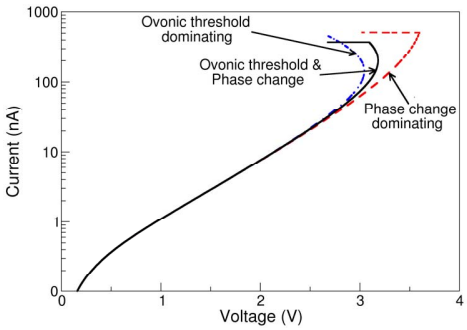

Figure 2 – Depending on the values of the thermal conductivity κ and of the energy relaxation time constant τ_R, different threshold events are evidenced in the I–V curve.

Thermal effects on transport equations

The fixed lattice temperature version of the model described in (10) proved its usefulness in investigating the smooth OTS event. Taking into account the energy dissipated to the lattice via Joule heating, the present enhanced version allows one to simulate also the much sharper MTS events due to the crystallization of the amorphous phase. According to the interplay between the energy relaxation time constant τ_R, the local electric field, and the transition temperature T_c of the chalcogenide material investigated (also depending on the thermal conductivity κ), three different conditions are found (Fig. 2). When the energy relaxation time is small carriers dissipate

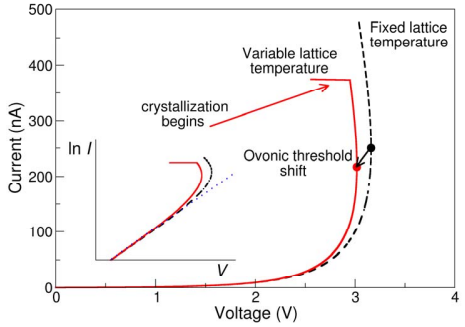

Figure 3 – Effects of the inclusion of the variable lattice temperature on the I–V characteristic. The inset shows data in a semilog scale to better highlight the increased slope of the super-exponential zone. The dotted line is an eye-guide of the pure exponential trend.

almost immediately the energy gained from the electric field via phonon scattering. The low mobility typical of the amorphous phase requires a relatively high electric field to sustain the current, so that Joule heating and MTS eventually dominate. On the other hand, when the energy transfer to the lattice is low (i.e., for a large value of the energy relaxation time) the energy gain due to the electric field enables the formation of a hot-carrier distribution. These conditions make the OTS event to happen. The third and last case is intermediate, so that hot carriers effects are limited. Depending on the transition temperature T_c, the MTS event may or may not be preceded by the OTS.

This last case is also analyzed in Fig. 3, where we compare two current-voltage (I–V) characteristics under constant and variable lattice temperature conditions. The inclusion of the local lattice temperature in the transport equations leads to three main consequences: 1) the presence of crystallization events; 2) a more pronounced super-exponential behavior preceding the switching events; and 3) the shift of the OTS condition ($dV/dI \approx 0$) towards lower voltages and currents.

Microscopic filamentary conduction

In Fig. 4 we show a computed I–V curve fitting experimental data from carbon-nanotube-contacted (CNT)-PCM devices, which represent the smallest PCM device volumes probed to date (4,11). Good agreement is obtained for both the subthreshold region and the predicted threshold. The simulation has been performed using the simulative domains with contacts as large as the active region (~10×10×40 nm^3). The following parameters have been used: $r_{cut} = 6$ nm, $\tau_0 = 49$ fs, $\tau_R = 5$ fs, $\ell / r_{cut} = 0.48$, $E_C = 0.48$ eV, $N = 48$.

The already mentioned common origin of OTS and MTS is further confirmed by the microscopic analysis of the simulation outcomes reported in Figs. 5 and 6. Fig. 5 describes the case of MTS only, whereas Fig. 6 has been obtained with a ~10× larger value of the energy relaxation time constant to include also the OTS. In both cases a preferential pathway for conduction is progressively created as long as the current is increased. In the so-called Ohmic and exponential conduction regimes, the current is initially almost evenly distributed on several paths; heat transfer to the lattice is negligible or moderate. When the exponential conduction turns into the super-exponential regime, the slight differences in the resistance among paths originating from the local disorder start to

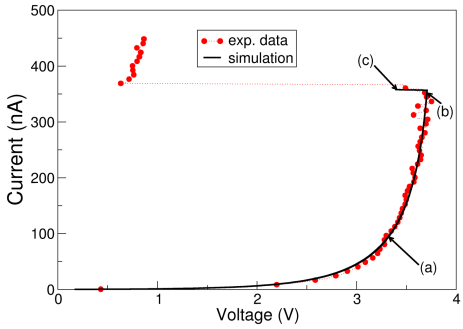

Figure 4 – I–V characteristic of a CNT test device (10×10×40 nm^3) compared to experimental data. Microscopic details on the conduction are given in Fig. 5 for the three points (a), (b), and (c).

play a major role. The current is then progressively gathered along a single path populated by slightly heated carriers. Thank to this current concentration, a significant Joule heating occurs so that crystalline segments between scattering centers populate the network. The presence of these low-resistance crystalline segments further boosts current concentration along the segments themselves, deactivating in practice any competing parallel route.

When OTS is present, as in Fig. 6, a strong self-increasing carrier heating near the OTS point enhances the creation of the preferential path. This, in turn, leads to two consequences: a) the existence of a negative-differential-resistance region, featuring hot carriers with high mobility, so that a large current can be sustained by a moderate electric field, and b) the inactivation of all parallel routes in the whole device when the crystallization of the preferential path occurs.

Threshold variability

Repeated experimental determinations of the I–V characteristic of a given device evidenced a certain degree of variability of the results (11). Neglecting the large fluctuations of the threshold condition in the very first fires, which are associated to structural rearrangements of the amorphous matrix created via sputtering and subsequently modified thermally, a residual variability is found also after the completion of such a burn-in (forming) process. These smaller fluctuations can be explained by considering that at each crystallization/re-amorphization cycle the positions of the defects within the amorphous matrix change. The random generation of the scattering centers in the present model can thus come into

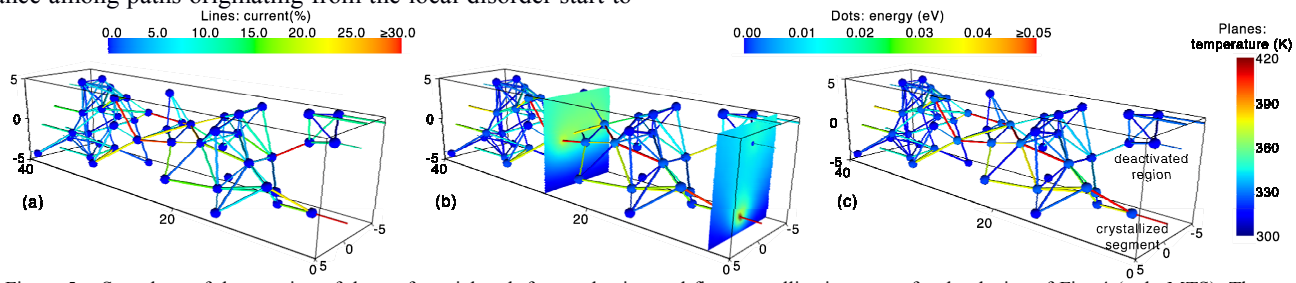

Figure 5 – Snapshots of the creation of the preferential path for conduction and first crystallization event for the device of Fig. 4 (only MTS). The two planes in panel (b) show the temperature profile on two selected cross-sections when crystallization begins.

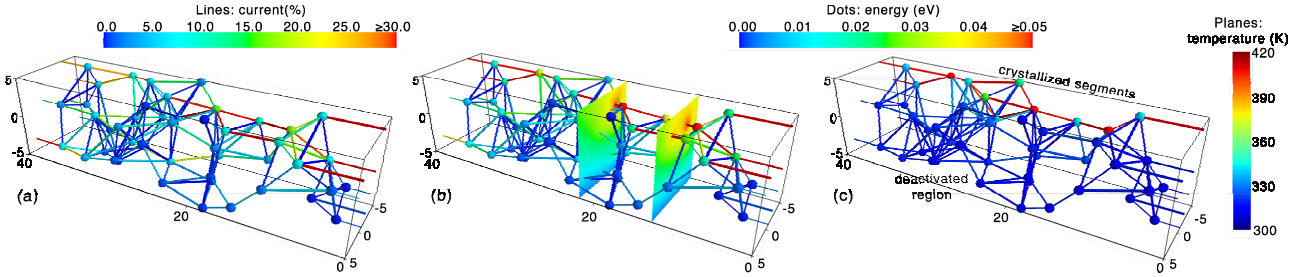

Figure 6 – Snapshots of the creation of the preferential path for conduction and first crystallization events when both OTS (panel a), and MTS (panel b) are present. Panel c) refers to a post MTS condition. Notice the high carrier temperature along the preferential path that was not present in Fig. 5.

hand in order to investigate these fluctuations.

To this aim, we have simulated batches of 100 geometrically identical samples, whose unique difference is the internal positions of the scattering centers. The analysis has been performed considering both traditional "mushroom-like" cells and the novel CNT-GST devices of (11), supposing an active chalcogenide volume of the same size ($10\times10\times40$ nm^3). In the latter case, in order to better represent the CNT contacts, we have also reduced the contact area to 2×2 nm^2. Results (Fig. 7) indicate that CNT-GST devices feature larger threshold voltages and currents, with a larger standard deviation for the latter. This can be ascribed to the larger power subtracted by the SiO_2 substrate onto which the device is built. We also notice that in "mushroom-like" devices the MTS was triggered by the OTS. No significant differences in the spread of the OTS data are found with respect to the MTS data shown in the two top panels of Fig 7.

Conclusion

A 3D random network model to investigate switching in nanoscale PCM devices has been derived within a full 3D trap-limited conduction framework. The model incorporates at a microscopic level both hot-carrier transport that could give origin to Ovonic switching, and Joule heating responsible for phase change and memory switching. The microscopic analysis of the model outcomes suggests that they share a common cause, i.e., the formation of preferential conductive paths; the interplay between the device parameters and the particular node configuration ultimately determines the switching types and in what order they occur. The model is also suitable for the study of the statistical dependence of device performances on to the intrinsic disorder of the amorphous bit.

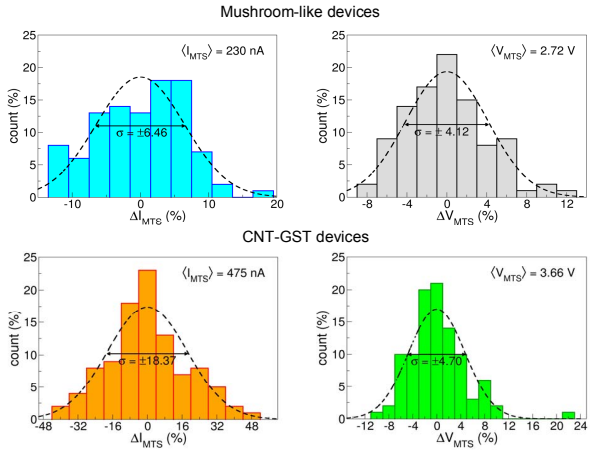

Figure 7 – Dispersion of the MTS point (V_{MTS}, I_{MTS}) for test devices having a similar active chalcogenide volume, but different layouts and contact areas. The two top panels refer to 100 "mushroom-like" devices, whereas the bottom panels to 100 CNT-GST devices with a 4 nm^2 contact area. Numerical results are presented as differences with respect to the average value of the distribution. Simulations have been performed with the interaction distance r_{cut} = 6.5 nm.

Acknowledgments

The authors gratefully acknowledge financial support from the Intel Corporation, the U.S. Office of Naval Research, and the U.S National Science Foundation.

References

(1) investors.micron.com/releasedetail.cfm?ReleaseID=692563
(2) H.-S. P.Wong *et al.*, "Phase change memory," *Proc. IEEE*, vol. 98, pp. 2201–2227, 2010.
(3) D. Kau *et al.*, "A stackable cross point phase change memory," in *IEDM Tech. Dig.*, pp. 617–620, 2009.
(4) F. Xiong, A.D. Liao, D. Estrada, and E. Pop, "Low-power switching of phase-change materials with carbon nanotube electrodes," *Science*, vol. 332, pp. 568–570, 2011.
(5) M. Nardone, M. Simon, I.V. Karpov, and V.G. Karpov, "Electrical conduction in chalcogenide glasses of phase change memory," *J. Appl. Phys.*, vol. 112, pp 071101 1-20, 2012.
(6) D. Ielmini and Y. Zhang, "Evidence for trap-limited transport in the subthreshold conduction regime of chalcogenide glasses," *Appl. Phys. Lett.*, vol. 90, pp. 192102 1-3, 2007.
(7) D. Ielmini, "Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses," *Phys. Rev. B*, vol. 78, pp. 035308 1-8, 2008.
(8) E. Piccinini *et al.*,"Hot-carrier trap-limited transport in switching chalcogenides," *J. Appl. Phys.*, vol 112 083722 1-11, 2012
(9) M. Rizzi and D. Ielmini, "Energy landscape model of conduction and switching in phase change memories" in *IEDM Tech. Dig.*, pp. 589–592,
(10) A. Cappelli *et al.*, "Conductive preferential paths of hot carriers in amorphous phase-change materials," *Appl. Phys. Lett.*, vol. 103, pp. 083503 1-4, 2013.
(11) F. Xiong *et al.*, "Self-aligned nanotube-nanowire phase change memory," *Nano Lett.*, vol.13, pp. 464-469, 2013.

5.3 Summary

In this Chapter a novel charge transport model was introduced, based on a Poole-like mechanism for carrier motion, which was employed to study conduction in amorphous chalcogenide devices, representing the latter as a random network of localized states.
A feature of the model is the possibility to approach the transport problem statistically, analyzing the IV curves obtained from a series of simulated devices, with the same macroscopic properties and parameters but different instances of the randomly generated network. This provides a way to address the technologically relevant issue of device variability and reliability.
The model allows to reproduce experimental IV curves from CNT-GST devices [129], providing estimates for the material-defining parameters (such as the characteristic times τ_0 and τ_R) consistent with values found in the previous literature [47, 45, 91, 50].
The microscopic analysis of the random network at several IV points correlates the Ovonic switching event to the formation of hot-electron preferential paths. They somehow remind of the μm-wide filaments predicted by Adler *et al.* [8], related to complex dynamics in the charge and nature of trap states. In the present model, however, the paths stem from non-equilibrium energy transport, and no specific quantitative claim is made as regards their physical size, since they represent preferred sequences of nodes the carriers move through. However, their existence in such scaled devices implies that their sizes must be estimated in the nanometer range.
An application of the model was explored, namely the analysis of the effect of the disorder degree in the network topology. Disorder is found to potentially impact both resistivity and threshold switching IV coordinates, a trend that can help interpreting a variety of phenomena related to structural rearrangements.
Finally, a thermal extension of the model was presented, allowing to both improve the description of the physics underlying the IV curves by accounting for enhanced thermal excitation in the subthreshold region, and predicting the onset of the crystallization event.

Chapter 6

PCM devices with graphene contacts

A major concern for phase-change memory applications is power-consumption [98, 4], especially as regards the very energy-demanding RESET operation. The energy needed to induce the phase transitions is proportional to the cell volume, thus a possible solution deals with reducing device dimensions, which would enable both energy-efficiency and a significant increase in device density in memory batches. For these reasons, a lot of efforts are being spent on testing the scalability of PCM-based devices, investigating new methods to confine the memory bit and exploring the ultimate limits of its performances.

Examples of this are the works by Xiong *et al.* [130, 129] and Liang *et al.* [68] with GST, employing Carbon NanoTubes (CNT) as contacts, in order to induce Joule heating and phase change in a very spatially limited active region of chalcogenide.
Figure 6.1 shows the RESET current levels achieved with CNT-contacted GST memory cells (CNT-GST) from [130] as a function of the electrode-chalcogenide contact area, in comparison with state-of-the-art devices. The concept devices show a $\sim$ 100x improvement on the best standard devices. Furthermore, the condition of approximately constant RESET current-density confirms that reducing the size of the active bit can be a way of improving the energy requirements, maybe exploiting novel nanostructured contacts.

A new possible and interesting candidate for this purpose is another carbon-based two-dimensional nano-material: graphene. Similarly to the case of CNTs, graphene can help defining a thin transport-active GST region. Furthermore, it features good conductance, high current capacity, transparency and flexibility [17], and recent fabrication

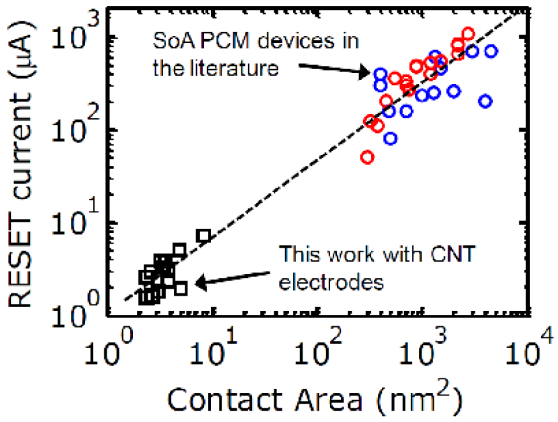

Figure 6.1: RESET current requirements for GST-based memory cells. State-of-the-art (SoA) industrial devices are compared with CNT-GST prototypes, with the latter demonstrating a $\sim$ 100x improvement with respect to the former. Figure taken from [130].

advances have allowed for the relatively inexpensive growth of high-quality graphene on a wafer scale [12].

Considering all of these properties combined, it is possible to envision the use of graphene as an interconnect material for transparent and/or wearable devices, where low-power GST memories could be integrated.

In this Chapter we describe the fabrication and electrical characterization of $Ge_2Sb_2Te_5$ devices contacted with graphene sheets employing two different approaches, namely nanometer-narrow graphene contacts with a GST-filled gap and micron-wide graphene sheets divided by a gap and bridged by a narrow GST line.

The topic presented in the following is the result of a collaboration with Prof. Pop's group, at the University of Illinois at Urbana-Champaign (UIUC) [16, 94].

6.1 Overview of transport properties of graphene interconnects

Since the work of Novoselov *et al.* [83], graphene has been the subject of a huge amount of research effort, on many different levels. Good and complete reviews on both general and specific topics can be found, e.g., in [82, 41, 37], but presenting the whole subject is well beyond the scope of the present work. Here we limit ourselves to making some practical considerations on the transport properties of graphene ribbons, and to analyzing the reasons why they could make good contacts for memory devices.

Graphene is a two-dimensional one-atom-thick hexagonal (honeycomb) lattice of sp^2-bonded carbon atoms [41]. In zero-temperature and no-defect conditions, an infinite graphene sheet exhibits a vanishing bandgap, corresponding to points in the Brouillin Zone (BZ) with a linear energy-momentum dispersion relation and consequently an effective electron mass equal to zero.
For most practical device purposes, the bandgap is found to be very small but non-vanishing, e.g., because of impurities, defects, bandstructure multi-layer effects (i.e. more than one graphene layer stacked, which give the material slightly different properties) or finite dimensions of the sheet or ribbon.
These band properties result in a distinctive dependence of graphene conductivity upon the position of the Fermi level E_F, yielding a recognizable V-shaped minimum when E_F is in the middle of the small (or zero) gap, and higher values when it goes deep into the conduction or valence band, making the material $n-$ or $p-$type accordingly. This behavior can be observed because of doping or through the field-effect induced via a gating voltage.

Behnam *et al.* [17] investigated the properties of Graphene NanoRibbons (GNR) with Au contacts, fabricated through Chemical Vapor Deposition (CVD) [126, 73], in a wide range of temperatures (1.9-900 K) and in both low- and high-electric field conditions, in order to assess their applicability as electrical interconnects. At room temperature, mobilities in the range 100 to 500 cm^2/Vs are found, and the material can sustain a peak current density $\sim 2 \times 10^9$ A/cm^2. These figures of merit, in addition to the above considerations about convenient integration in flexible/transparent systems and GST nano-confinement, motivate the investigation of graphene for contacting memory devices.

Electrical modeling parameters A detailed microscopic description of electron transport in graphene involves a sophisticated theoretical framework [82], of great interest by itself and currently subject of research effort, at various regimes of temperature, electric field [35], and structural conditions (disorder, defects, dimensions) [84, 82, 15]. However, as it will be discussed later, experimental uncertainties make it unpractical in our case to account for such level of detail, neither we think it can severely hinder our analysis at the present stage.
In what follows, graphene will be modeled as a series resistance with appropriate circuit parameters. The devices studied here are not gated, so that the carrier concentration does not vary due to a gating voltage. Nonetheless, it is safe to assume that some degree of uncontrolled chemical doping is present. Furthermore, interactions with the substrate

(silicon dioxide, as explained later) or structural defects and imperfections can similarly affect carrier concentration in graphene. However, the total carrier density is known to vary only within a factor of some units as a consequence of that.

For these reasons, the following parametrization will be adopted: a sheet resistance $R_s^g \sim 3000\ \Omega/\square$ will be assumed for graphene, as an average measure of its circuit electrical properties. The contact resistance R_c^{g-m} of graphene (g) at the interface with metal leads (m), gold and nickel in our case (m =Au,Ni), will be estimated through the relation $R_c^{g-m} = \rho_c^{g-m}/W_r$, with W_r the width of the graphene ribbon (and thus the metal-graphene contact region) and ρ_c^{g-m} the corresponding contact resistivity per unit width. We will assume $\rho_c^{g-\mathrm{Au}} = 0.25\ \Omega$cm and $\rho_c^{g-\mathrm{Ni}} = 1.0\ \Omega$cm. These values are

Table 6.1: Electrical parameters of graphene

Quantity	Symbol	Value
graphene sheet resistance	R_{sh}^g	$3000\ \Omega/\square$
g-Au contact resistivity per width	$\rho_c^{g-\mathrm{Au}}$	$0.25\ \Omega$cm
g-Ni contact resistivity per width	$\rho_c^{g-\mathrm{Ni}}$	$1.0\ \Omega$cm

summarized in Table 6.1. Values for these quantities were obtained from graphene-only reference devices, fabricated on purpose to estimate these parameters in the same experimental conditions as the final GST devices.

6.2 Device design and fabrication

In order to investigate the effect of different possible architectures, two geometries were explored.

In the first concept, GNR were used (ribbon width $W_r \sim$ tens of nanometers). The bit area was defined by opening a transverse GNR-wide gap at one point along the length of the ribbon, and subsequently depositing GST over a rectangular region covering also the gap region, filling it with GST at one spot and making it the active bit. We will refer to these devices as GNR-GST. See Figure 6.2(a).

In the second, somewhat complementary case, Graphene MicroRibbons (GμR, $W_r \sim$ micrometers) were used. Once again, a transverse gap was cut, but now a nm-wide GST line (GST NanoRibbon, GSTNR, width W_{GSTNR}) was deposited, bridging the two parts of the separated GμR, and specifically defining the GST filled point of the gap as the

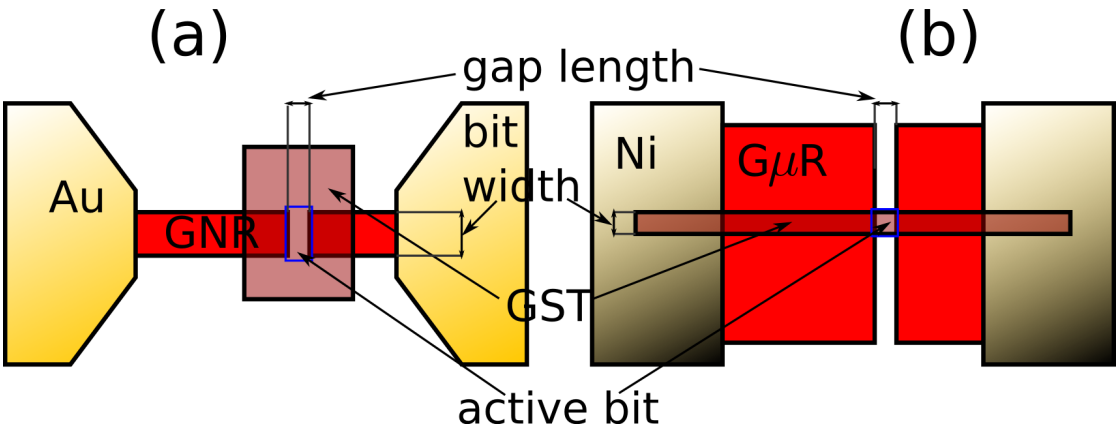

Figure 6.2: Scheme of the designs for GNR-GST (a) and GSTNR (b) devices. See text.

active bit. See Figure 6.2(b).
In both cases, the intersection of the transverse gap (with gap length L_G) and the deposited GST layer allows one to obtain a confined nm-wide and -long area. This, combined with the small thickness t of the GST, ~ 10 nm, provides a small volume of material to SET and RESET during memory operations, with volume $\approx W \times L_G \times t$; W is the effective width of the active region, given by W_r in the GNR-GST case, and by W_{GSTNR} for the micro-ribbon structure.

The fabrication of GNR-GST and GSTNR devices is a complex process involving multiple lithography (both photolithography and e-beam lithography) and deposition phases. Figure 6.3 summarizes the common steps between the two device typologies and the differences in the two processes.
First, a graphene layer is fabricated (steps 1 to 6): graphene is deposited by CVD on both sides of a copper layer (1), then a Polymethylmethacrylate (PMMA) layer is deposited onto one side of the graphene-Cu-graphene sandwich, and the graphene layer not in contact with PMMA is removed via O_2 plasma etching (2). The copper is chemically etched away using $FeCl_3$ (3,4) and the PMMA-graphene structure is transfered onto a Si++/SiO_2 substrate, with graphene directly touching the 90-nm-thick silicon oxide (5). Finally, PMMA is chemically removed (using acetone or dichloromethane), and large probing metallic pads are patterned onto the graphene by combining photolithography and e-beam evaporation (6).
Regarding GNR-GST, the metal chosen for the large pads is gold over titanium, Ti/Au. E-beam lithography is employed to define smaller contacts with Pd/Au (7a) and the nanoribbons dimensions and gap (8a). Plasma etching is then used to remove excess

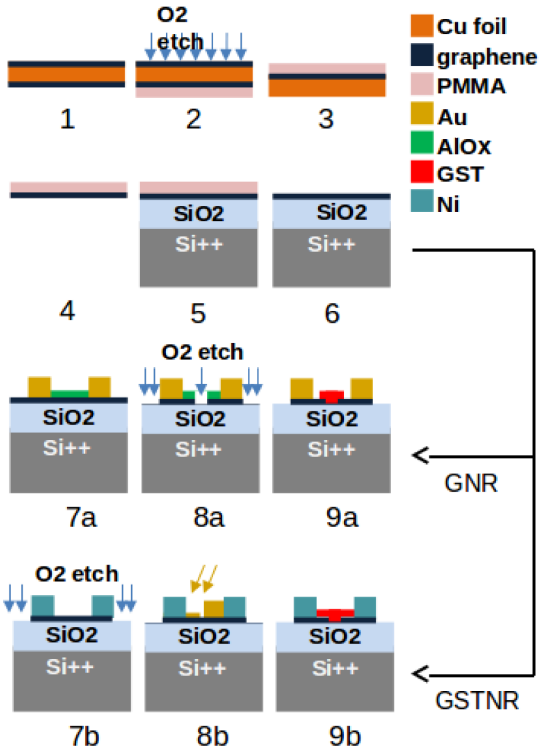

Figure 6.3: Fabrication phases for GNR-GST and GSTNR devices, see text [18].

graphene outside the device region; an aluminum oxide layer protects the graphene ribbon during etching and is subsequently removed. At last, a $\sim$ 10-nm-thick GST rectangular "window" is deposited in the gap area using e-beam lithography to define the region for deposition (9a), and a protective SiO_2 layer is evaporated on top of the complete structure. Typical dimensions are: gap length $L_G = 30 - 100$ nm, GNR length $L_r = 0.5 - 1.0$ μm and GNR widths $W_r = 30 - 400$ nm, the latter also determining the bit width.

As for the GSTNR, starting from the large contacts (nickel over titanium in this case, Ti/Ni), graphene is patterned into microribbons (2-10 μm-wide) by photolithography in the region between the contacts, and it is removed from everywhere else (7b). A gold-based shadow evaporation technique [77] is combined with O_2 plasma etching to define a nm-long gap in the GμR (8b). After this, a GST nanowire is defined employing e-beam lithography and deposited across the gap (9b). Finally, a protective 10-nm-thick silicon oxide layer is evaporated on top.

Figure 6.4 shows micrographs of the results of the fabrication procedures described

above. Panel 6.4a shows the optical image of one row of large pads, with the green circles

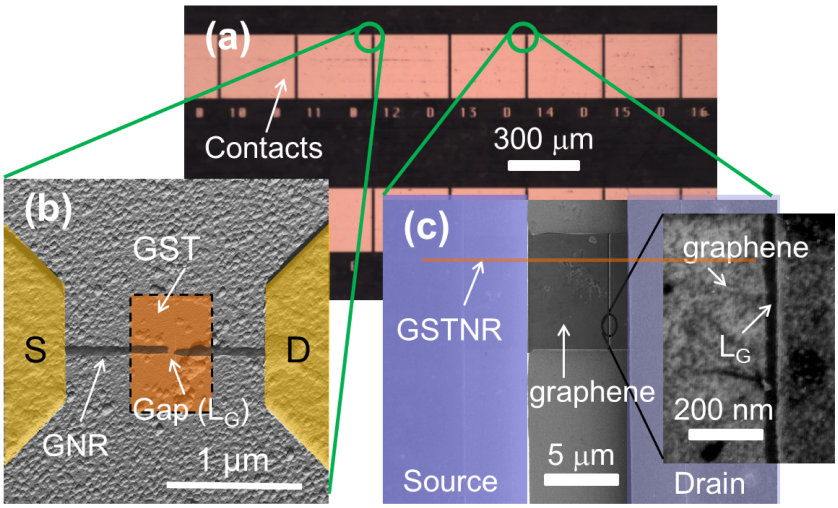

Figure 6.4: Images of device structures: (a) optical image of the macroscopic contact pads on Si++/SiO_2; (b) Atomic Force Microscope image of a GNR-GST device, with the GST window deposited above the ribbons shown in brown; (c) Scanning Electron Microscope image of a GSTNR device, with the GST (brown) line crossing the gap in the graphene microribbon. The gap region is shown with further magnification [18].

indicating the regions where the nanodevices are located. Figures 6.4b and 6.4c show images of typical geometries for both designs.

Figure 6.4b represents a GNR-GST device. The yellow regions are nano-patterned areas of the border of the large pads, playing the roles of Source (S) and Drain (D) contacts. The nanoribbon portions (dark horizontal line) extend from below the contacts, and are divided by a gap near the middle point. The brown area represents the GST-covered region.

Figure 6.4c shows the GSTNR design. The large blue areas on the sides represent the S and D pads. The graphene microribbon is visible in the upper-center part of the picture, with the vertical gap cut near its right side. A further magnification of a GST-free segment of the gap is shown. The chalcogenide nanowire is marked by the brown horizontal line, going from one pad to the other, crossing the microribbon and, of course, the gap. The GST bridge is not shorting the two contacts directly, since the GST is originally deposited in the amorphous phase, and it reasonably makes the most resistive path in all cases but for the crossing of the gap, which makes the gap itself the only significantly GST-active spot.

The two structure typologies are convenient means for two different purposes: the

GNR-GST devices focus on the investigation of the scalability potential of GST with ultimate scaling of the graphene electrodes, in this new planar, graphene-based configuration. The GSTNR, instead, allow for a more practical extraction of the relevant descriptive parameters of the material, due to the better defined device structure and subsequent lower uncertainty.

With the purpose of comparing the performances of graphene-contacted devices with those of more traditional designs, e.g., with metal contacts, a batch of "control" devices was fabricated in a way analogous to the GNR-GST case, with Au nanoribbons instead of graphene.

6.3 Device characterization

We performed the electrical characterization with the experimental setup represented in the scheme of Figure 6.5. Device chips were located in a Janis VTPS vacuum probesta-

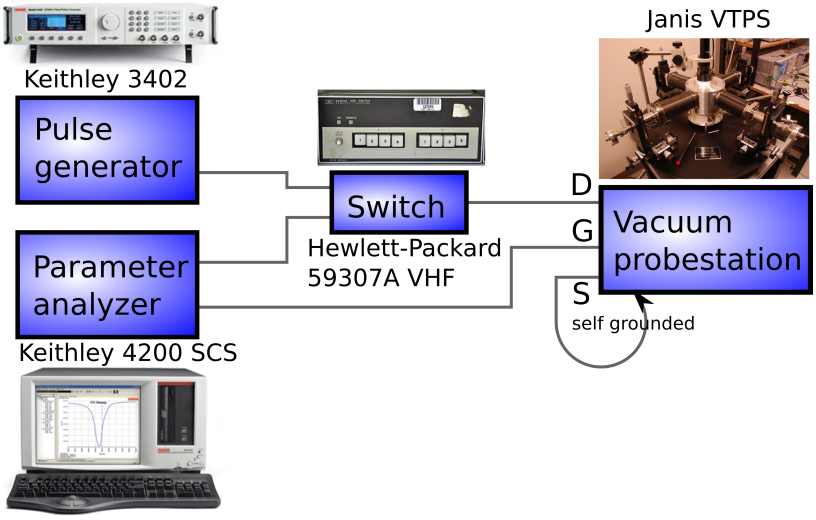

Figure 6.5: Experimental setup for the electrical characterization of graphene-contacted GST memory devices. See text.

tion, in which a turbo pump was used to maintain a vacuum of 10^{-5}–10^{-4} torr, to prevent oxidation or contamination from atmospheric gases. The probe tips were connected to couples of metal pads to close the circuit containing the device between them. The source tip (S) was shorted to the external metallic body of the chamber, in order to set its potential to ground. The drain tip (D) was connected to a Hewlett-Packard 59307A VHF

switch box, and from there to a Keithley 4200 SCS Semiconductor Parameter Analyzer (SPA) tool. The stage on which the chip was placed (G) was also connected to the SPA, with its potential set to 0, and monitored during measurements to detect possible leakage currents through the substrate of chip. The switch box allowed one to change the drain connection from the SPA to a Keithley 3402 Pulse Generator Unit (PGU), employed for generating the programming pulses. The PGU could produce voltage pulses with amplitude in the range $0-20$ V and with variable duration time t_d, and rising (t_r) and falling (t_f) edge times, with minimum available durations of the order of some nanoseconds. RESET operations were performed with short applied pulses, with nominal $t_r = t_f \sim 2.5$ ns and $t_d \sim 50$ ns, and amplitudes often above 5 V, depending on the dimensions of the devices. SET programming was usually obtained during current-driven IV measurements, since crystallization was obtained once the SET current I_{th} was reached. This allowed to test the crystallization capabilities of the device while also measuring their threshold current and voltage.

6.3.1 IV characteristics

To test the electrical behavior of the devices, current-driven IV measurements were run, with freshly-fabricated devices starting in the as-deposited amorphous phase. Figure 6.6 shows results from a clearly switching GNR-GST device, which was SET and RESET several times to test the repeatability of its behavior. The device under test had $L_G \sim 80$ nm and bit width $W \sim 30$ nm. Qualitatively similar behaviors were found for GSTNR, too. The red curves represent the low-resistance crystalline state, which was

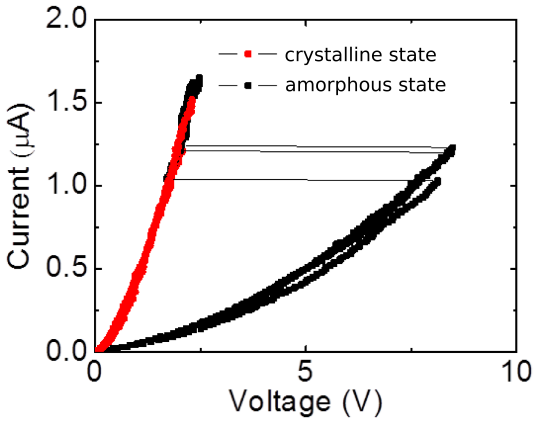

Figure 6.6: Typical IV curve for a GNR-GST device with $L_G \sim 80$ nm and $W \sim 30$ nm. A very sharp snapback is found in the current-driven measurement for the amorphous state (black line). The subsequent measure (red line) confirms the happened crystallization [18].

found very similar after several SET-RESET. The black curves come from the amorphous high-resistance state, which showed a slightly larger variability, but exhibited quite similar resistance and switching after each test.
Some characteristic curves were found to be a bit more noisy than Figure 6.6, sometimes featuring deviations from the ideal exponential behavior at mid-range electric voltages, and/or small partial phase-changes. This kind of behavior is mostly ascribed to imperfections along the fabrication process, and can be due to issues such as a non-uniform distribution of GST, or to defects at the GST-graphene interface, related to deposition-induced damages of the graphene structure.

6.3.2 Analysis of the switching behavior

As anticipated, the IV curves did not always show the same exact shape, even more so in the vicinity of the threshold point. In some cases, in near-threshold conditions the characteristic presented a very high-slope portion, occasionally with several small snap-backs and snap-forwards, which could make it difficult to define a precise single switching point. Here we describe the criteria we adopted to extract the threshold behavior parameters from the IV graphs.
We considered the SET switching as the event where the voltage was seen to decrease abruptly and significantly. This also corresponds to the system going into a new conduction regime (i.e. a different IV curve) and a new structural phase, which could be recovered in a subsequent measurement. Thus, this threshold corresponds to a SET operation. This description allows to define a SET threshold current I_{th} quite naturally as the prescribed value of I at which this snap-back occurs.
As regards the measured threshold voltage V_{th}^{m}, it is defined as the highest value of voltage reached in the IV before the snap-back described above occurs. A scheme to clarify these definitions is shown in Figure 6.7. The threshold voltage V_{th}^{m} belongs to the same point to which I_{th} belongs for situations such as the one in Figure 6.6, also sketched in curve (2) of Figure 6.7. A different case can also occur, namely the sudden snap-back happens only after a smooth decrease in the voltage, which in these circumstances can be identified as the Ovonic Threshold Switching. In this occurrence no phase change is implied (not yet, at least) and, in voltage-driven measurements, an applied bias above such value would lead to a much higher current, belonging to the high-conductivity state of the amorphous material. The case of OTS≠SET switching is represented by curve (1) in Figure 6.7, where two distinct threshold points are shown, one associated to the threshold current $I_{th}^{(1)}$, the other to the threshold voltage $V_{th}^{m(1)}$.

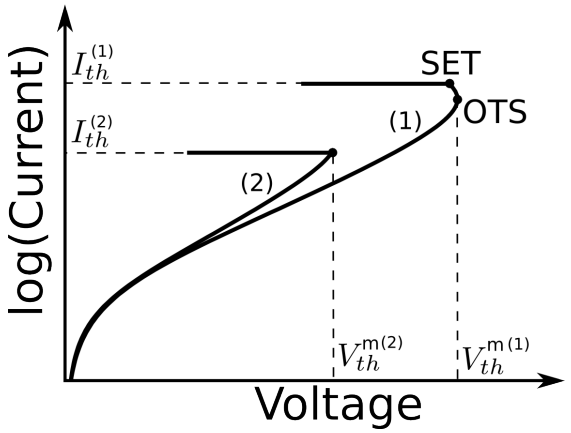

Figure 6.7: Scheme representing the possible switching situations, see text.

The intrinsic V_{th}, used for plotted data, is actually given by $V_{th} = V_{th}^{m} - IR_{not-bit}$, where I is the prescribed current and $R_{not-bit}$ is the estimated series resistance not originating from the GST-active region itself. In this case, it consists of the graphene ribbon, its contact resistance at the graphene-metal interface (both when applicable), and the external circuit (including the metal pads). The latter is assumed to have a negligible resistance. The graphene resistance was estimated on the basis of the dimensions of the ribbon, according to $R = R_{sh}^{g} \times L_r/W_r$, with L_r and W_r the length and width of the ribbon, respectively. R_{sh}^{g} is the sheet resistance of graphene; its value and the formulas and parameters for the metal-graphene contact resistances are those reported in section 6.1. We note in passing that this way of calculating V_{th} somehow assumes that the graphene-GST interface is considered as part of the region on which V_{th} drops, which is reasonable by an operative point of view.

From the curves collected over measurements of all three kinds of devices (GST-GNR, GSTNR, control), the threshold current densities j_{th} and threshold fields F_{th} were extracted, assuming $t = 10$ nm and the bit width W and gap length L_G equal to their design values. Figure 6.8 shows the results of such analysis. Threshold fields, computed as $F_{th} = V_{th}/L_G$ are found in the range 0.5-10 MV/cm, with a normal-like distribution (in lin-log scale) that gathers the majority of values between 0.7 and 3 MV/cm. Generally, the control devices showed slightly lower critical fields, but all species varied approximately in the same range.
The observed fields are compatible with the value reported for other lateral devices, such as $F_{th} = 0.56$ MV/cm, measured in GST bridge devices with TiN contacts [61], and $F_{th} \sim 1.0$ MV/cm in carbon-nanotube-contacted GST memory cells [130]. The

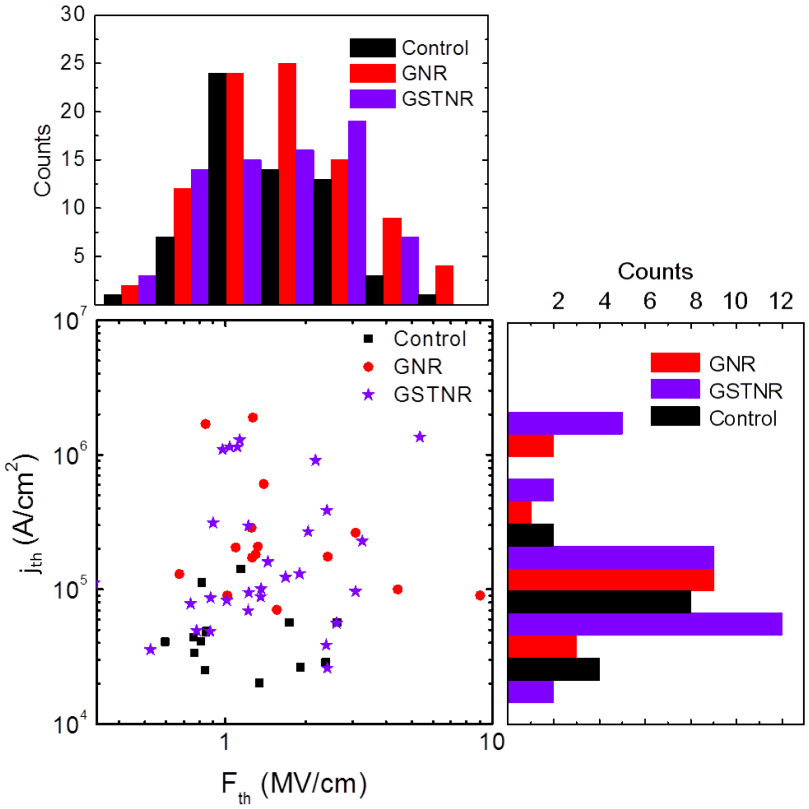

Figure 6.8: Plot of the threshold current-density j_{th} against the estimated threshold field F_{th}. The right- and top-most graphs represent the frequencies of occurrence of the two quantities, respectively [18].

discrepancies can be due to a number of factors: the L_G used for calculations in the graphene case can be different from the design value, yielding an incorrect estimation of the critical field, or other fabrication imperfections can alter the electrical behavior of the device with respect to the ideal case. Also, the field reported in [61] considers the GST-only voltage, subtracting the potential drop due to the GST-contact interface, differently from the case at hand. Accounting for that contribution would likely yield a better comparison.

We found threshold current densities $j_{th} \approx I_{th}/(Wt)$ in the 2×10^4–3×10^6A/cm^2 range. Again, control devices were observed to have slightly lower threshold current densities with respect to graphene-involving structures.

A relation typically observed in both industrial [46] and carbon-contacted devices

[130, 129] is the correlation of the low-field resistance vs the threshold voltage. To investigate it, a $R - V_{th}$ plot is usually drawn, expecting a rough proportionality between the two, based on their ideal common dependence on the length of the device (or more precisely on the length of its amorphized region in the RESET state).
Figure 6.9 shows such plot for the GSTNR devices, for both the SET and RESET resistances, in the hypothesis of full crystallization and amorphization of the active region, respectively. Plotted R is given by the bit-only contribution, calculated as the measured values minus the graphene and graphene-metal contact contributions. Furthermore, since the resistance depends on the different widths of the active bits, it is reported normalized to the bit width W. In the ideal case, under this assumption, the measurements of

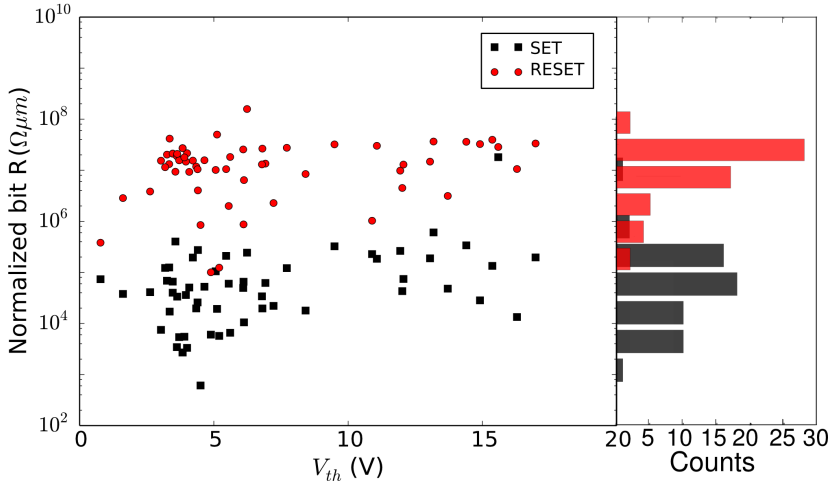

Figure 6.9: Resistances of the GSTNR devices for both RESET (red circles) and SET (black squares) states, normalized to the device width W (i.e. the quantity $R' = R \times W$ is shown), plotted against the threshold voltage V_{th}. The histograms on the side panel show the occurrencies of the resistance values [18].

resistance in both states, as well as V_{th}, should approximately be proportional to the bit length L_G, yielding a linear correlation between the two.
A significant dispersion was observed in the data from the graphene-contacted bits. However, a relationship can be found in a statistical sense, i.e. the lowest values of R_{RESET} and R_{SET} are more likely found in correspondence of a low V_{th}. The side histograms of Figure 6.9 depict the occurrences of R values.

6.3.3 Performance analysis

Since the scaling study of GST devices is one of the main reasons for this investigation, we measured the RESET current densities j_{RESET} needed to fully amorphize the GNR-GST and GSTNR bits. They were obtained by applying a RESET pulse with an amplitude V_{RESET} to a cell in the crystalline state, the minimum one such that full amorphization was obtained. The low-field resistance R_{SET} of the SET state prior to amorphization, assumed to be independent from the electric field, was used to estimate $j_{RESET} \approx V_{RESET}/(R_{SET}Wt)$.
The results are reported in the histogram of Figure 6.10. The GST-GNR structures show typically lower values of j_{RESET}, with respect to the GSTNR case. The lower ends of the two ranges of values are very similar, while the upper value appears definitely higher for GSTNR. The range of values is $\sim 10^5$–5.6×10^7 A/cm^2 and is roughly compatible

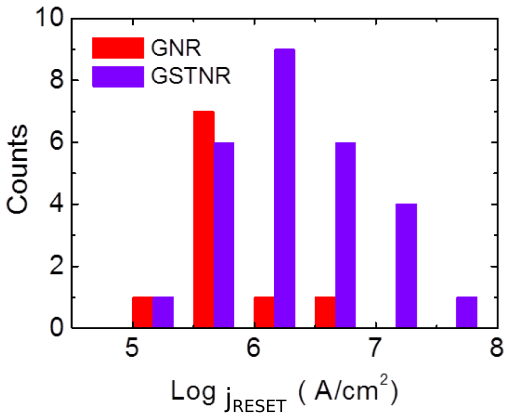

Figure 6.10: Distribution of RESET current values, for GNR-GST and GSTNR devices [18].

with values from self-aligned carbon-nanotube-contacted devices [129], having $j_{RESET} \approx 6.4 \times 10^7$ A/cm^2. As mentioned, these calculations assume that the current spreads over the whole bit cross-section, due to the lack of further microscopic information, and it is possible that the current density is somewhat higher than reported, due to the current flowing through a smaller effective cross-section. In that case, these estimates should be considered as a lower bound to the real value. Anyway, it seems quite reasonable that a great fraction of the active bit is actually involved in the full RESET operation, and that current passes through most of its area.

Another important figure of merit is the ratio between the low-field resistances of the RESET and SET states, i.e. $r_{R/S} = R_{RESET}/R_{SET}$. It gives a reasonable idea of

how easily distinguishable the two states are under READ conditions, even in presence of noise or other disturbing factors. Furthermore, it can hint at the potential use of devices for multi-bit programming [87, 88, 23]. A wide range of resistances between the full-amorphous and full-crystalline values is not a sufficient property, since stability of the intermediate states over time can be the biggest issue; however, it is an important pre-requisite to even think of such multi-level application.

Figure 6.11 shows the R_{RESET} plotted against the R_{SET} of GNR-GST, GSTNR and control devices with metal contacts, with both quantities normalized to the bit width after removal of the graphene and graphene-metal contributions. The Figure shows dashed lines representing sets of (R_{SET}, R_{RESET}) points having the same $r_{R/S}$, to guide the eye and define regions corresponding to significantly different ratios.

While both control and GNR-GST devices are found mostly in the 10^0–10^3 region (the GNR-GST mainly between 10^2 and 10^3), the GSTNR devices span the entire 10^0–10^4 range, with very few cases below 10, the majority of data in the 10^2–10^3 region and a significant amount of cases above $r_{R/S} = 10^3$, up to very close to the 10^4 line. Although RESET pulses were applied in order to obtain full amorphization, i.e. to probe the maximum resistivity of the devices, and similarly the SET switching were used to test the fully crystalline state, it is possible that some of the operations instead brought the device to intermediate amorphous/crystalline states. This would make the measured SET resistances higher than the best achievable, and conversely the RESET ones lower than the highest possible. Thus, the results in Figure 6.11 can be regarded as lower

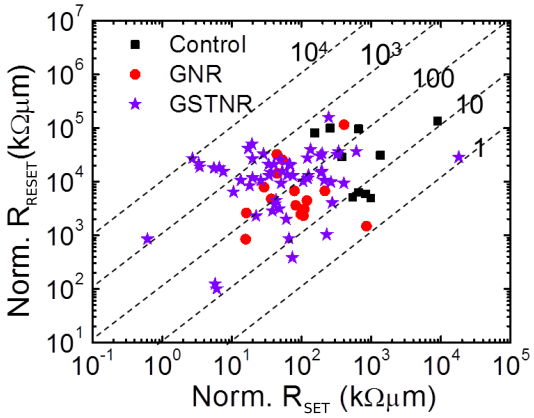

Figure 6.11: Correlation plot of device low-bias resistances R in the RESET (y-axis) and SET (x-axis) states, both normalized to the device width W (i.e. the quantity $R' = R \times W$ is shown). The dashed lines are guides for the eye, representing the sets of points having the same R_{RESET}/R_{SET} ratio, as indicated near each line. Both GNR-GST and GSTNR data are reported, together with control data from metal-contacted memory cells [18].

bounds for the $r_{R/S}$ performances of the devices, since any lower-than-reported R_{SET} or higher-than-reported R_{RESET} would push the data even further toward the top-left corner of the graph, i.e. toward better ratios.
However, the best achieved values $\lesssim 10^4$ are probably a good estimate of the best possible performances for the GSTNR, practically hitting the intrinsic limits of resistivity contrast in the material. The resistivities $\rho_a \approx 10^2$ Ωcm and $\rho_c \approx 10^{-2}$ Ωcm were assumed for the amorphous and crystalline phases, respectively [129], hence $\rho_a/\rho_c \sim 10^4$. Also, these ratios compare quite well with what is observed for self-aligned CNT-GST devices [129], where the typical $r_{R/S}$ are ~ 452 (versus similar average values with graphene) and some devices showed $\sim 2 \times 10^3$ value.

It is known that a reversible memory switching can be induced in SiO_2 [134, 133, 135], in particular exploiting nano-confinement of electric fields, as it is done here, with carbon nanotubes and graphene. Since silicon dioxide is involved in the fabrication of the devices described in this Chapter (substrate and protective capping layer), it is wise to reject its possible role in switching only after careful consideration.
Batches of devices in the various structures were fabricated skipping the GST deposition phase, so that graphene ribbons were obtained, laying on an SiO_2 substrate and also with a gap filled in with silicon dioxide. These devices did not exhibit any switching behavior. Furthermore, the typical characteristic values observed in this work for GST devices are close to what actually expected for GST, e.g. current levels ($\sim \mu A$ at switching), $r_{R/S}$ (a maximum of $\sim 10^4$, versus ratios of $\sim 10^5$ or more observed in SiO_2 [134]), and threshold fields around a few MV/cm (while one more order of magnitude would be expected for oxide switching).

Finally, we attempted to estimate a measure of the quality of the contacts at the GST-graphene interface, which is related to the concept of contact resistance. The contact resistance R_c^{m-x} is due to the effect that the interface between materials m (usually the conductive lead) and x (the material "under test") and their immediately adjacent regions have on the current flow. This interface behavior is often very conveniently modeled through the specific contact resistivity ρ_c^{m-x} [108] at the interface, since ρ_c^{m-x} is a contact-area-independent quantity that generalizes the concept of resistance per unit area, while including the geometry-dominated effect of current transferring from one material to the other. Such quantity also allows for comparisons between different device architectures. In the following analysis, the m role is played by graphene ($m = g$), and x is the amorphous ($x = a$) or crystalline ($x = c$) GST.
The contact resistance in a memory device can play an important role in the programming

operations, since it is desirable that Joule power is mostly dissipated in the inner part of the bit, instead of the interfaces, to make the programming more efficient and avoid power waste. This behavior can be partially hindered by too high contact resistances, which can absorb a relevant fraction of the pulse power.
The estimation of the contact resistance was performed by modeling our devices as *contact front resistance test structures* (see Chapter 3 of [108]), as shown in the scheme in Figure 6.12. The contact resistivity ρ_c^{g-x} is calculated as the solution to the equation

$$R_c^{g-x} = \frac{\sqrt{R_{sh}^x \rho_c^{g-x}}}{Z} \coth\left(\frac{L_c\sqrt{R_{sh}^x}}{\sqrt{\rho_c^{g-x}}}\right) = \frac{\rho_c^{g-x}}{L_T Z} \coth\left(\frac{L_c}{L_T}\right) \tag{6.1}$$

where R_c^{g-x} is the measured contact resistance (see below), Z is the width of the contact region between the two materials and L_c is the length of such region (see Figure 6.12a). $L_T = \sqrt{\rho_c^{g-x}/R_{sh}^x}$ is the transfer length parameter, defined as the position inside the contact region, in the transport direction, at which the electric potential has come to be a fraction $1/e$ of its value at the beginning of such region. It represents the region through which most of the current passes at the contact.

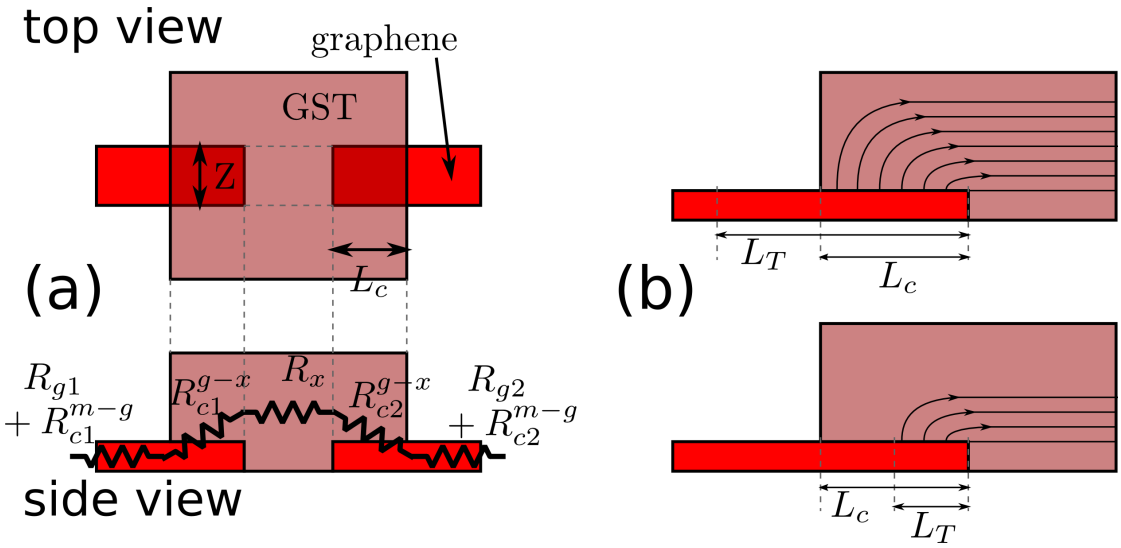

Figure 6.12: Schematic representations of the contact region parametrization and electrical modeling from the top and side views (a), and pictorial description of the concept of transfer length L_T in the limiting case of L_T much greater or much smaller than L_c, upper and lower pictures, respectively (b).

To gain more insight into the physical meaning of the transfer length concept, it is possible to look at the two extreme cases of the relation between L_c and L_T: when

$L_c \ll L_T$ (Figure 6.12b, upper picture), Eq. (6.1) reduces to $R_c^{g-x} = \rho_c^{g-x}/(ZL_c)$, which means that, since the overlap length between the graphene and the GST is much smaller than the length L_T in which the current transfer would typically happen, all the current must make the transition within a length L_c. This is formally equivalent to actually considering a uniform transfer resistivity ρ_c^{g-x}, over an effective area $Z \times L_c$. In the opposite situation, $L_c \gg L_T$ (Figure 6.12b, lower picture), the formula becomes $R_c^{g-x} = \rho_c^{g-x}/(ZL_T)$, analogously to the previous case. This indicates that in this case the effective area becomes equal to $Z \times L_T$, i.e. smaller than the available one ($= Z \times L_c$), since "all" the current is able to cross the interface in a smaller region. It must be noted that, in our modeling, we assumed for simplicity that $Z = W$, thus neglecting the effect of lateral current spread.
For practical purposes, the two limit cases above are actually frequently used, since they are already quite accurate for $L_c \leq 0.5 \times L_T$ and $L_c \geq 1.5 \times L_T$, respectively [108]. However, in the following, Eq. (6.1) was solved numerically, with no such assumptions.

Calculations were performed estimating the contact resistance R_c^{g-x} according to the equivalent circuit shown in the lower picture of Figure 6.12a: the total measured resistance R was considered as the resistance series of the inner region of the GST bit R_x, the graphene-related resistance due to both the ribbon parts themselves R_{g1}, R_{g2} and their contact resistances at the metal-graphene interfaces R_{c1}^{m-g}, R_{c2}^{m-g} (assumed to be known and calculated as explained earlier), and finally the graphene-GST contact resistances R_{c1}^{g-x}, R_{c2}^{g-x}. Symbols with different numerical indeces are used for corresponding quantities on the two sides of the bit, but symmetry considerations were used whenever possible to treat these pairs of values as equal, thus simplifying calculations. In the symmetric case, $R_c^{g-x} = \left[R - (R_{g1} + R_{c1}^{m-g}) - (R_{g2} + R_{c2}^{m-g}) - R_x\right]/2 = [R - R_{not-bit} - R_x]/2$. R_c^{g-x} was then used for the numerical solution of Eq. (6.1). The sheet resistance of GST was assumed to be $R_{sh}^a = 10^8\ \Omega/\square$ and $R_{sh}^c = 10^4\ \Omega/\square$ for the amorphous and crystalline states, respectively, based on the resistivity values assumed in this work for GST resistivities [129].
Figure 6.13 shows the statistics of the results of such contact resistivity calculations for the RESET (left panel) and SET (right panel) states of GSTNR, GNR-GST and metal-contacted control devices. The graphene-contacted structures exhibit a quite wide range of values, for contact resistivities in both phases: for the GSTNR, the values lie in the interval 3×10^{-5}–$3 \times 10^{-2}\ \Omega/\text{cm}^2$ for the amorphous GST-graphene interface, and 10^{-6}–$3 \times 10^{-3}\ \Omega/\text{cm}^2$ for the crystalline GST-graphene interface. The GNR-GST have shown similar values, with a slightly higher upper bound for the RESET state and a slightly higher lower bound for the SET state. The control contact resistivities are found

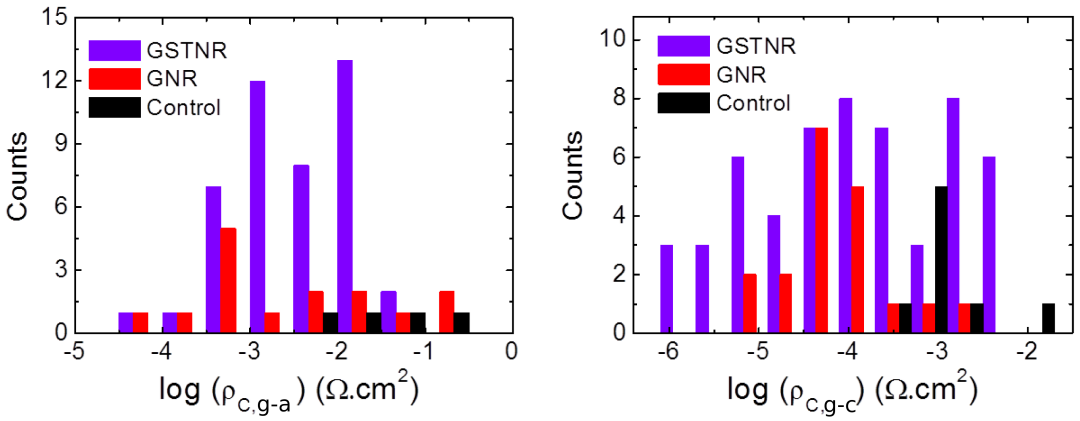

Figure 6.13: Distribution of estimated contact resistivities, between GST (both amorphous and crystalline, left and right panel, respectively) and graphene in the two GNR-GST and GSTNR designs, compared with control metal-GST devices [18].

in the upper region of both these intervals, with $\rho_c^{g-a} \sim 10^{-2}$–$0.3\ \Omega/\text{cm}^2$ for RESET and $\rho_c^{g-c} \sim 3 \times 10^{-4}$–$2 \times 10^{-2}\ \Omega/\text{cm}^2$ for SET. The graphene cases compare quite favorably with the metal control group: although the wide spread of values means that maybe the fabrication can be improved, the average and best values with graphene are both found significantly better than the correspondent parameters in metal-contacted devices.
The figures we find are also comparable with measurements from the literature, where more sophisticated structures specifically fabricated for contact resistivity investigation were employed: Circular Transmission Line Method measurements [108] on TiW-contacted devices report $\rho_c^{TiW-a} \approx 7 \times 10^{-4}\ \Omega/\text{cm}^2$ and $\rho_c^{TiW-c} \approx 5 \times 10^{-6}\ \Omega/\text{cm}^2$ (SET measurements at 200 °C) [104], while Cross Bridge Kelvin Resistance measurements [108] yielded $\rho_c^{TiW-a} \approx 9.5 \times 10^{-2}\ \Omega/\text{cm}^2$, and $\rho_c^{TiW-c} \approx 4 \times 10^{-6}\ \Omega/\text{cm}^2$ and $\rho_c^{TiW-c} \approx 2 \times 10^{-7}\ \Omega/\text{cm}^2$ at 175 and 350 °C, respectively [106]. Similar values are found also in thin films of doped Sb_2-Te with TiW contacts [105], i.e. $\sim 10^{-7}$–10^{-6} and $\sim 10^{-4}$–$10^{-2}\ \Omega/\text{cm}^2$ for the SET and RESET states, respectively, with the increasing values of the latter due to increasing thickness of the layer in specific experiments.
While the values in the crystalline devices are comparable with but hardly match the best values found in previous works, the amorphous devices often show ρ_c^{g-a} in good agreement with the literature, with a significant number of best values representing an improvement on already existing devices.

6.4 Summary

In this Chapter we described the fabrication and characterization of the first CVD graphene-contacted GST memory prototypes. Two structures were employed, namely graphene nanoribbons with a central gap filled by depositing a GST layer, and micron-wide ribbons, also with a transverse gap, bridged by a GST nanowire. The two structures aimed at analyzing the scaling properties of graphene contacts: the former focused on the scalability potential offered by the ultra-thin and -narrow GNR, while the latter, more stable, was chosen for better parameter extraction.
The devices showed varied behaviors, possibly due to the challenging fabrication process, but the best samples featured threshold voltages as low as $\sim$3 V, low programming currents ($\lesssim$ 1 μA SET, $<$ 10 μA RESET) and $>$ 100 typical RESET/SET resistance ratio, with best cases approaching the 10^4 limit value. Contact resistances were found comparable to what reported in the literature for conventional contact materials.
Despite the fabrication-related issues, these devices showed promising scaling outlooks, and represent a step forward along the way to assess both the intrinsic and feasibility limits of ultra-scaled PCM devices.

Chapter 7

Conclusions

Aggressive scaling has been, and still is, one of the most defining trends of the electronic industry. The area of memory devices is no exception, and the scaling problem is an even more central topic for phase-change memories: they represent a very promising solution for the next generation of memory devices, but their relatively high power requirements make it mandatory to seek ways to shrink the devices, so that SET/RESET operations on lower bit volumes are less energy demanding.
Here, we will summarize the research activity described in this work and aimed at investigating the conduction and switching properties of nanoscale phase-change memory devices. Then, we will review the main findings of such activity and, finally, hint at some possible directions for future research.

The present thesis revolved around the study of charge transport and switching properties of nano-scale phase change memory devices, both by theoretical modeling and experimental electrical characterization.
On the modeling side, the main challenges consisted in the understanding of the amorphous phase, featuring a rich physics and a variety of intriguing phenomena, with important technological consequences, particularly as device dimensions are reduced.
At first, a quasi-1D hydrodynamic-like continuous model was devised to capture both the subthreshold and the switching behaviors of an amorphous chalcogenide memory cell. A physical trap-limited hot-electron picture was taken as a starting point to elaborate a full simulation scheme, featuring a proper analysis of the carrier statistics and distribution along the device, consistent with electrostatics and accounting for carrier diffusion. The same model was applied to the analysis of how different density-of-state distributions in the amorphous material can affect conduction.

A further level of analysis moved to consider a 3D domain, where the continuum approximation was abandoned in favor of a network approach, where discrete nodes represented carrier-hosting centers. The effect of different configurations of such centers could be studied by means of their randomized generation, to mimic the intrinsic variability of the amorphous phase. The same model formed the basis of an analysis of the impact of thermal phenomena on charge transport, including the onset of crystallization.
On the experimental side, the performances of ultra-confined bits of phase-change materials were studied by using graphene ribbons as device contacts. Several different test structures allowed to both probe the potential and limits of the material in very confined geometries, and estimate some figures of merit.

The above mentioned efforts allowed to gain some insight into the physics ruling transport in chalcogenides, and the potential and limits of their exploitation as ultra-scaled memory devices.
The continuous quasi-1D theory was able to correctly reproduce the typical IV behavior of present-day industrial memory cells. The fitting agreement was also quantitatively good. The proper analysis of electrostatics and diffusive effects led to different (but comparable) characteristic process times with respect to literature data.

The 3D network theory allowed to tackle the transport problem from a new perspective, by applying the concepts of the quasi-1D theory to a 3D randomized system. While a good fit of ultra-scaled device data was again found, it was also possible to analyze how the electrical properties depend on the stochastic nature of the system, and which model parameters regulate this dependence. Although no *a priori* prediction can be made at the present stage of development for any sort of device design, it is possible to calibrate the network model on the observed experimental device-to-device variations for each design, in order to characterize its degree of reliability.
The 3D model also enabled a microscopic analysis of carriers at the network level, suggesting that the onset of Ovonic switching in amorphous chalcogenides is associated to the formation of preferential paths for high-energy electrons.
Finally, the thermal extension of the analysis correctly reproduced the crystallization behavior observed in real devices, and provided a qualitative assessment of how the material parameters can interact to determine which switching (Ovonic or crystallization) is the triggering one for a given device.

The experimental investigation of ultra-scaled graphene-contacted devices provided a useful insight in what the challenges and rewards of aggressive scaling can be. Even with conservative analysis assumptions, several figures of merit of these devices appeared

quite promising, such as a high resistance ratio between RESET and SET state and good current density values, which couple with the particular advantages of employing graphene interconnects, such as transparency and flexibility. On the other hand, the scattering of the collected data observed in device performance statistics highlighted how the complex fabrication procedure likely has room for improvement, that can be explored in further investigations.

Despite the results described here, several developments can be foreseen, both as extensions of the past activity and as parallel investigations of complementary aspects of the topic. Let us finally consider some of them.
The theoretical tools developed in this thesis can be adapted to a variety of situations. The hydrodynamic model was developed in the assumption of a uniform density of states in a given range of energy inside the bandgap. However, this simplifying hypothesis can be relaxed to include more complex and realistic energy structures, as derived from experiments. This would probably prevent some calculation advantages, but it still could provide relevant insight into the device physics.
Furthermore, the steady-state condition can be substituted by a full time-dependent analysis, accounting for all transient phenomena that were not considered at the present stage, and can play a significant role when crystallization sets in.

On the network simulation side, several paths can be explored: since traps can be arbitrarily distributed in a given domain, the model lends itself to investigating a variety of possible geometries, and their impact on conduction and reliability.
Several interesting physical phenomena are associated to the idea of a self re-adjusting amorphous structure. In this regard, it is important to cite the very technologically-relevant resistance drift which implies an increase in the RESET resistance as time passes. It is possible to envision a simulation framework (such as, e.g., Kinetic Monte Carlo) allowing to study the dynamics of the network nodes over time, including their motion in the computational domain, appearance or disappearance, which the network model can immediately relate to the corresponding evolution of resistance over time. This would provide useful insight into the drift problem.
Clearly, a straightforward and very important follow-up to this work would be the validation of the statistical aspect of the model with a significant amount of data, particularily as regards the device-to-device variability.
The picture at the basis of the simulation framework is intrinsically semiclassical. However, the correctness of such starting point must be questioned when device dimensions become such that quantum effects can heavily modify device operations. It would be

worth to make an effort toward a quantum treatment of the transport problem, both to assess the consequences of device shrinking, and to test the quantum mechanical correctness of the many simplifying assumptions made at present at the modeling level.

Acknowledgements

There is a number of people to thank for their direct or indirect contribution to this thesis work. They made it possible by providing guidance, scientific help, moral support or a combination of the three. I will try to group them according to some criteria (geographically-wise, relationship-wise), as far as possible.

Italian side, UniMoRe First of all, I thank my advisor, Prof. Carlo Jacoboni, for his guidance during my PhD path. He has always shown a helpful attitude, and the willingness to listen to my ideas or suggestions and provide the appropriate feedback, i.e. encouragement for the good ones and the necessary red-light for the crazier, unviable ones. I also thank him for driving me to cultivate my own scientific critical sense, and to "walk with my own legs" every day a little more.
I owe very much to Dr. Enrico Piccinini, with whom much of the "heavy-lifting" for this PhD was planned. Due to his meticulous nature and his extremely germane observations, discussions with him resulted in either revealing debugging sessions (sometimes in the coding sense, sometimes in the conceptual one) or precious brainstormings. I hope he will consider my use of the word "germane" in this case to be germane. I don't want to jeopardize the formal correctness of these Acknowledgements.
A grateful mention goes also to Prof. Rossella Brunetti from the Monte Carlo Laboratory and Prof. Massimo Rudan from the University of Bologna, who were always willing to provide their scientific insight and share their experience. Not less importantly, I thank them for financial support for conferences and missions, within Contract No. 3477131/2011 from Intel Corporation, whose support is sincerely acknowledged.
I thank Dr. Fabrizio Buscemi and Prof. Paolo Bordone for advice and discussion on the themes of Physics and Life in general, and their always helpful and positive attitude. Thanks to Dr. Andrea Bertoni and Prof. Luca Zanni for their technical help and scientific insight. Their suggestions and opinion in several circumstances ended up saving me time, providing a solid background to mere conjectures and/or dissipating doubts.

I want to thank the colleagues and collaborators from Bologna, particularly Dr. Fabio Giovanardi and Giuliano Marcolini, and Dr. Ilaria Imperiale, whose experience and advice made my life much easier during the preparation for my UIUC stay.
Thanks to the several office mates who shared work days with me in the Monte Carlo Lab, sometimes for a B.S. thesis lifespan, sometimes for an M.S. or a Ph.D. one, particularly Claudia Benedetti and Andrea Beggi among them. Thanks to you all, for patiently bearing my occasional questions, doubts and erratic thoughts, and for addressing them with your feedback.
Thanks to the mates of many lunches together, among the other Francesca Gherpelli, Elena Cipressi, Annamaria Boni, Paolo Restuccia, Luca Bursi, Alfredo Bonini, Simone Ierinò, Andrea Schiavi, and to all those who felt at home inside Lab. 1 (obviously, the same set of people very well knows the only, true name of such place). Thank you 1000! times. The exclamation mark is exactly where it should be.
Thanks to all the colleagues, classmates and friends from the Physics Department, with a mention for the historical ones: Giulio Guzzinati, Mirko Franchini, Alberto Artioli, Francesco Casolari, Stefano Gavioli, Gianluca Santoni, Lucio Anderlini, Sabrina Magnani, Ugo Tricoli, Riccardo Sangiorgio, Davide Meglioli, Gioia Cattini. Things have changed quite a bit in the past years, and we ended up being spread here and there around the world, but it's always a pleasure to meet all of you whenever possible, be it for a reunion lunch, a random encounter or just a couple of lines written over Skype. You guys are great and I wish you all the best, whatever path you are pursuing.
Thanks to my "students" in the Foundations of Theoretical Physics course. I hope I didn't load them with an unjustified amount of entropy, and that they'll find some use or value in at least a small fraction of what we talked about together. It was a privilege to know you, and I guess I got more than I gave. Thanks also to Prof. Carlo Maria Bertoni, for providing such opportunity.
Thanks to all of the PhD students of the Physics Department in Modena, belonging to all Cycles, particularly Ilaria Siloi for moral support in multiple situations, and for the randomly distributed but very effective "PhD Anonymous"-like meetings in front of the coffee vending machines.
Thanks to all UniMoRe professors, researchers and technicians for their help, suggestions and guidance in these years.

US side, UIUC I am very grateful to Prof. Eric Pop for welcoming me in his group at the University of Illinois and providing me with the means I needed for being trained

and then dive into an activity which was new to me. My experience at UIUC taught me a lot from several, quite orthogonal points of view, and he is definitely one of the individuals I must thank the most for making it possible.
I need to thank Dr. Ashkan Behnam. He was my tutor "on-the-field" during my stay at UIUC. He always provided a helping hand and proved to be an outstanding researcher and professional, an enjoyable office mate and a good friend.
Likewise, I need to thank Sungduk Hong. He was my first contact with the PopLab. Since the very first moment he always showed a unique willingness to help, and always behaved as a real friend. I thank him for his support in so many circumstances (one above all, The Great Move Of The 100-Tons-Television), for both the scientific and fun chats during lunches at the Beckman Institute, and for making me know his wonderful family.
Thanks to all other members of the PopLab as well, particularly Feng Xiong, Kyle Grosse and the other Memory Subgroup guys, for their help, technical assistance and feedback. Their enthusiasm and expertise have been a source of inspiration.
Thanks to Massimiliano Bianchi, for the [SitCom+JunkFood+Conan O'Brian] evenings, for the raids at KFC, for the perfectly timed and carefully engineered missions to the laundry building, for Jack O'Led and more, and, in a broad sense, for his Outstanding Achievements in the Field of Roommateness.

Italian side, home Finally, I must thank the people that basically were always there. Thanks to the many Modena and Bologna friends, met at very different stages of my life. They behaved the same way as always, even after I disappeared, quite literally in some cases, once in a while. A special mention to the GC group.
Thanks to my grandmother Lidia, my aunt Silvia, and Sauro, for general (but very effective and welcome, nonetheless) support.
Thanks to Anna and Franco, for making everything possible in the first place.

Sara Thank you. For everything. Thank you.

Bibliography

[1] http://investors.micron.com/releasedetail.cfm?ReleaseID=692563. p. 29

[2] See Supplementary Material at http://dx.doi.org/10.1063/1.4761997 for an animation showing the evolution of f(z), t(z), and n(z) as a function of the current. p. 49

[3] See Supplementary Material at http://dx.doi.org/10.1063/1.4819097 for three animations describing the microscopic evolution of the network as a function of the prescribed current. p. 70

[4] International Technology Roadmap for Semiconductors. http://www.itrs.net/Links/2011ITRS/Home2011.htm. p. 47 , p. 95

[5] V. N. Abakumov, V. I. Perel, and I. N. Yassievich. *Nonradiative recombination in semiconductors*. Elsevier, 1991. p. 43

[6] N. Kh. Abrikosov and G. Danilova-Dobryakova. An investigation of the structural diagram of Sb_2Te_3-GeTe. *Izvestiya Akademii Nauk SSSR, Neorganicheskie Materialy*, 1:204–207, 1965. p. 12

[7] D. Adler, H. K. Henisch, and N. Mott. The mechanism of threshold switching in amorphous alloys. *Reviews of Modern Physics*, 50:209–220, 1978. p. 32 , p. 33 , p. 47

[8] D. Adler, M. S. Shur, M. Silver, and S. R. Ovshinsky. Threshold switching in chalcogenide-glass thin films. *Journal of Applied Physics*, 51(6):3289–3309, 1980. p. 33 , p. 34 , p. 47 , p. 93

[9] V. Agafonov, N. Rodier, R. Ceolin, R. Bellissent, C. Bergman, and J. P. Gaspard. Structure of Sb_2Te. *Acta Crystallographica Section C: Crystal Structure Communications*, 47(6):1141–1143, 1991. p. 12

[10] J. Akola and R. O. Jones. Structural phase transitions on the nanoscale: The crucial pattern in the phase-change materials $Ge_2Sb_2Te_5$ and GeTe. *Physical Review B*, 76(23):235201, 2007. p. 15 , p. 16 , p. 49

[11] P. W. Anderson. Model for the electronic structure of amorphous semiconductors.

Physical Review Letters, 34(15):953, 1975. p. 43

[12] S. Bae et al. Roll-to-roll production of 30-inch graphene films for transparent electrodes. *Nature Nanotechnology*, 5(8):574–578, 2010. p. 96

[13] S. A. Baily and D. Emin. Transport properties of amorphous antimony telluride. *Physical Review B*, 73(16):165211, 2006. p. 41

[14] S. A. Baily, D. Emin, and H. Li. Hall mobility of amorphous $Ge_2Sb_2Te_5$. *Solid State Communications*, 139(4):161–164, 2006. p. 23 , p. 24 , p. 41

[15] V. Barone, P. Hod, and G. E. Scuseria. Electronic structure and stability of semiconducting graphene nanoribbons. *Nano Letters*, 6(12):2748–2754, 2006. p. 97

[16] A. Behnam et al. Phase change memory with graphene ribbon electrodes. Oral Presentation at the MRS Spring Meeting, 2013. p. 96

[17] A. Behnam, A. S. Lyons, M.-H. Bae, E. K. Chow, S. Islam, C. M. Neumann, and E. Pop. Transport in nanoribbon interconnects obtained from graphene grown by chemical vapor deposition. *Nano Letters*, 12(9):4424–4430, 2012. p. 95 , p. 97

[18] A. Behnam, F. Xiong, A. Cappelli, N. Wang, S. Hong, E. Carrion, Y. Dai, A. S. Lyons, E. K. Chow, E. Piccinini, C. Jacoboni, and E. Pop. To be submitted for publication. p. v , p. vi , p. 100 , p. 101 , p. 103 , p. 106 , p. 107 , p. 108 , p. 109 , p. 113

[19] G. Betti Beneventi, L. Guarino, M. Ferro, and P. Fantini. Three-dimensional Poole-Frenkel analytical model for carrier transport in amorphous chalcogenides. *Journal of Applied Physics*, 113(4):044506, 2013. p. 38 , p. 69

[20] N. A. Bogoslovskij and K. D. Tsendin. Electronic–thermal switching and memory in chalcogenide glassy semiconductors. *Journal of Non-Crystalline Solids*, 357(3):992 – 995, 2011. p. 43

[21] N. A. Bogoslovskiy and K. D. Tsendin. Physics of switching and memory effects in chalcogenide glassy semiconductors. *Semiconductors*, 46(5):559–590, 2012. p. v , p. 16 , p. 29 , p. 33 , p. 42 , p. 43 , p. 44 , p. 47

[22] N. A. Bogoslowsky and K. D. Tsendin. Nonlinearity of current-voltage characteristics of chalcogenide glassy semiconductors, caused by multiphonon tunnel ionization of negative-U-centers. *Semiconductors*, 43(10):1338–1342, 2009. p. 43

[23] S. Braga, A. Cabrini, and G. Torelli. Experimental analysis of partial-set state stability in phase-change memories. *IEEE Transactions on Electron Devices*, 58(2):517–522, 2011. p. 109

[24] F. Buscemi, E. Piccinini, R. Brunetti, M. Rudan, and C. Jacoboni. Monte Carlo simulation of charge transport in amorphous chalcogenides. *Journal of Applied Physics*, 106(10):103706, 2009. p. 47

[25] F. Buscemi, E. Piccinini, M. Rudan, R. Brunetti, and C. Jacoboni. Diffusion and high-frequency noise of electrons in amorphous semiconductors at low electric fields. *Fluctuation and Noise Letters*, 11(03):1242004, 2012. p. 47

[26] A. Calderoni, M. Ferro, D. Ielmini, and P. Fantini. A unified hopping model for subthreshold current of phase-change memories in amorphous state. *IEEE Electron Device Letters*, 31(9):1023–1025, 2010. p. iv , p. 36 , p. 38

[27] A. Cappelli, E. Piccinini, F. Xiong, A. Behnam, R. Brunetti, M. Rudan, E. Pop, and C. Jacoboni. Conductive preferential paths of hot carriers in amorphous phase-change materials. *Applied Physics Letters*, 103(8):083503, 2013. p. 70 , p. 86 , p. 88

[28] S. Caravati, M. Bernasconi, T. D. Kühne, M. Krack, and M. Parrinello. Coexistence of tetrahedral- and octahedral-like sites in amorphous phase change materials. *Applied Physics Letters*, 91(17):171906, 2007. p. 15

[29] S. Caravati, M. Bernasconi, T. D. Kühne, M. Krack, and M. Parrinello. First-principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects. *Journal of Physics: Condensed Matter*, 21(25):255501, 2009. p. 49

[30] M. Cardona and R. Merlin. *Light scattering in solids IX*. Springer, 2007. p. 14

[31] M. Cassinerio, N. Ciocchini, and D. Ielmini. Evidence for electrically induced drift of threshold voltage in $Ge_2Sb_2Te_5$. *Applied Physics Letters*, 103(2):023502, 2013. p. 39

[32] E. Cho, J. Im, C. Park, W.-J. Son, D.-H. Kim, H. Horii, J. Ihm, and S. Han. Atomic and electronic structures of amorphous $Ge_2Sb_2Te_5$: melt-quenched versus ideal glasses. *Journal of Physics: Condensed Matter*, 22(20):205504, 2010. p. iii , p. 14 , p. 49

[33] N. Ciocchini, M. Cassinerio, D. Fugazza, and D. Ielmini. Modeling of threshold-voltage drift in phase-change memory devices. *IEEE Transactions on Electron Devices*, 59(11):3084–3090, 2012. p. 39

[34] N. Derhacobian, S. C Hollmer, N. Gilbert, and M. N. Kozicki. Power and energy perspectives of nonvolatile memory technologies. *Proceedings of the IEEE*, 98(2):283–298, 2010. p. 47

[35] V. E. Dorgan, M.-H. Bae, and E. Pop. Mobility and saturation velocity in graphene on SiO_2. *Applied Physics Letters*, 97(8):082112, 2010. p. 97

[36] D. Emin. Current-driven threshold switching of a small polaron semiconductor to a metastable conductor. *Physical Review B*, 74(3):035206, 2006. p. 41

[37] A. C. Ferrari et al. Raman spectrum of graphene and graphene layers. *Physical*

Review Letters, 97(18):187401, 2006. p. 96

[38] I. Friedrich, V. Weidenhof, W. Njoroge, P. Franz, and M. Wuttig. Structural transformations of GeSbTe films studied by electrical resistance measurements. *Journal of Applied Physics*, 87:4130–4134, 2000. p. iv , p. 19 , p. 20 , p. 21

[39] H. Fritzsche. Physics of instabilities in amorphous semiconductors. *IBM Journal of Research and Development*, 13(5):515–521, 1969. p. 67

[40] H. Fritzsche. Why are chalcogenide glasses the materials of choice for Ovonic switching devices? *Journal of Physics and Chemistry of Solids*, 68(5):878–882, 2007. p. 67

[41] A. K. Geim and K. S. Novoselov. The rise of graphene. *Nature Materials*, 6(3):183–191, 2007. p. 96 , p. 97

[42] J. Hegedüs and S. R. Elliott. Microscopic origin of the fast crystallization ability of Ge-Sb-Te phase-change memory materials. *Nature Materials*, 7(5):399–405, 2008. p. 15

[43] H. K. Henisch. Amorphous-semiconductor switching. *Scientific American*, 221:30–41, 1969. p. 33

[44] B. Huang and J. Robertson. Bonding origin of optical contrast in phase-change memory materials. *Physical Review B*, 81(8):081204, 2010. p. 16

[45] D. Ielmini. Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses. *Physical Review B*, 78:035308, 2008. p. 3 , p. 24 , p. 35 , p. 38 , p. 39 , p. 47 , p. 48 , p. 69 , p. 93

[46] D. Ielmini, A. L. Lacaita, and 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. p. 32 , p. 106

[47] D. Ielmini and Y. Zhang. Analytical model for subthreshold conduction and threshold switching in chalcogenide-based memory devices. *Journal of Applied Physics*, 102(5):054517, 2007. p. iv , p. 35 , p. 36 , p. 37 , p. 39 , p. 47 , p. 48 , p. 69 , p. 93

[48] D. Ielmini and Y. Zhang. Evidence for trap-limited transport in the subthreshold conduction regime of chalcogenide glasses. *Applied Physics Letters*, 90(19):192102, 2007. p. iv , p. 3 , p. 23 , p. 30 , p. 36 , p. 37 , p. 69

[49] C. Jacoboni. *Theory of Electron Transport in Semiconductors: A Pathway from Elementary Physics to Nonequilibrium Green Functions*. Springer, 2010. p. 48

[50] C. Jacoboni, E. Piccinini, F. Buscemi, and A. Cappelli. Hot-electron conduction in Ovonic materials. *Solid-State Electronics*, 84:90–95, 2013. p. 49 , p. 69 , p. 93

[51] V. G. Karpov, Y. A. Kryukov, I. V. Karpov, and M. Mitra. Field-induced nucleation in phase change memory. *Physical Review B*, 78(5):052201, 2008. p. 39 , p.

40

[52] V. G. Karpov, Y. A. Kryukov, S. D. Savransky, and I. V. Karpov. Nucleation switching in phase change memory. *Applied Physics Letters*, 90(12):123504, 2007. p. iv , p. 39 , p. 40 , p. 47

[53] T. Kato and K. Tanaka. Electronic properties of amorphous and crystalline $Ge_2Sb_2Te_5$ films. *Japanese Journal of Applied Physics*, 44:7340, 2005. p. 18 , p. 22 , p. 23 , p. 24

[54] D-C. Kau et al. A stackable cross point phase change memory. In *IEDM Technical Digest*, pages 617–620, 2009. p. 47 , p. 69

[55] J.-J. Kim, K. Kobayashi, E. Ikenaga, M. Kobata, S. Ueda, T. Matsunaga, K. Kifune, R. Kojima, and N. Yamada. Electronic structure of amorphous and crystalline $(GeTe)_{1-x}(Sb_2Te_3)_x$ investigated using hard x-ray photoemission spectroscopy. *Physical Review B*, 76(11):115124, 2007. p. 23

[56] K. Kohary and C. D. Wright. Electric field induced crystallization in phase-change materials for memory applications. *Applied Physics Letters*, 98(22):223102–223102, 2011. p. 18

[57] K. Kohary and C. D. Wright. Modelling the phase-transition in phase-change materials. *Physica Status Solidi B*, pages 944–948, 2013. p. 18

[58] A. V. Kolobov, P. Fons, A. I. Frenkel, A. L. Ankudinov, J. Tominaga, and T. Uruga. Understanding the phase-change mechanism of rewritable optical media. *Nature Materials*, 3(10):703–708, 2004. p. iii , p. 12 , p. 13 , p. 14 , p. 15 , p. 16

[59] B. T. Kolomiets and N. A. Goryunova. Svoistva i struktura troinykh poluprovodnikovykh sistem. 1. Elektricheskie svoistva i struktura nekotorykh materialov v sisteme Ti-Sb-Se. *Zhurnal Tekhnicheskoi Fiziki*, 25:984, 1955. p. 47

[60] M. Krbal, A. V. Kolobov, P. Fons, J. Tominaga, S. R. Elliott, J. Hegedus, and T. Uruga. Intrinsic complexity of the melt-quenched amorphous $Ge_2Sb_2Te_5$ memory alloy. *Physical Review B*, 83(5):054203, 2011. p. 16

[61] D. Krebs, S. Raoux, C. T. Rettner, G. W. Burr, M. Salinga, and M. Wuttig. Threshold field of phase change memory materials measured using phase change bridge devices. *Applied Physics Letters*, 95(8):082101, 2009. p. 31 , p. 105 , p. 106

[62] D. Kuzum, R. G. D. Jeyasingh, B. Lee, and H.-S. P. Wong. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Letters*, 12(5):2179–2186, 2011. p. 3

[63] A. L. Lacaita, A. Redaelli, D. Ielmini, F. Pellizzer, A. Pirovano, A. Benvenuti, and R. Bez. Electrothermal and phase-change dynamics in chalcogenide-based memories. In *IEDM Technical Digest*, pages 911–914, 2004. p. 31

[64] S. Lai and T. Lowrey. A 180 nm nonvolatile memory cell element technology for stand alone and embedded applications. In *IEDM Technical Digest*, pages 803–806, 2001. p. 48

[65] S. Lavizzari, D. Sharma, and D. Ielmini. Threshold-switching delay controlled by $1/f$ current fluctuations in phase–change memory devices. *IEEE Transactions on Electron Devices*, 57(5):1047–1054, 2010. p. 32 , p. 39

[66] B.-S. Lee. *Optical and electronic properties, nanoscale structural order, and transformation kinetics of phase change materials.* PhD thesis, 2006. p. 23

[67] B.-S. Lee, J. R. Abelson, S. G. Bishop, D.-H. Kang, B. Cheong, and 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. *Journal of Applied Physics*, 97(9):093509, 2005. p. iv , p. 18 , p. 19 , p. 20 , p. 22 , p. 23

[68] J. Liang, R. G. D. Jeyasingh, H.-Y. Chen, and 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. p. 95

[69] G. Lucovsky and R. M. White. Effects of resonance bonding on the properties of crystalline and amorphous semiconductors. *Physical Review B*, 8(2):660, 1973. p. 16

[70] H.-K. Lyeo, D. G. Cahill, B.-S. Lee, J. R. Abelson, M.-H. Kwon, K.-B. Kim, S. G. Bishop, and B. Cheong. Thermal conductivity of phase-change material $Ge_2Sb_2Te_5$. *Applied Physics Letters*, 89(15):151904, 2006. p. 19 , p. 21 , p. 22 , p. 23

[71] T. Matsunaga and N. Yamada. Structural investigation of $GeSb_2Te_4$: A high-speed phase-change material. *Physical Review B*, 69:104111, 2004. p. 11

[72] T. Matsunaga, N. Yamada, and Y. Kubota. Structures of stable and metastable $Ge_2Sb_2Te_5$, an intermetallic compound in GeTe-Sb_2Te_3 pseudobinary systems. *Acta Crystallographica Section B*, 60(6):685–691, 2004. p. 11

[73] C. Mattevi, H. Kim, and M. Chhowalla. A review of chemical vapour deposition of graphene on copper. *Journal of Materials Chemistry*, 21(10):3324–3334, 2011. p. 97

[74] A. Mendoza-Galvan and J. Gonzalez-Hernandez. Drude-like behavior of Ge:Sb:Te alloys in the infrared. *Journal of Applied Physics*, 87(2):760–765, 2000. p. 19

[75] N. F. Mott. Conduction and switching in non-crystalline materials. *Contemporary Physics*, 10(2):125–138, 1969. p. 33

[76] N. F. Mott and E. A. Davis. *Electronic Processes in Non-Crystalline Materials.* Clarendon, Oxford, 1979. p. 18 , p. 48 , p. 69

[77] Y. Naitoh, M. Horikawa, H. Abe, and T. Shimizu. Resistance switch employing a simple metal nanogap junction. *Nanotechnology*, 17(22):5669, 2006. p. 100

[78] S.-W. Nam, H.-S. Chung, Y. C. Lo, L. Qi, J. Li, Y. Lu, A. T. C. Johnson, Y. Jung, P. Nukala, and R. Agarwal. Electrical wind force–driven and dislocation-templated amorphization in phase-change nanowires. *Science*, 336(6088):1561–1566, 2012. p. 17

[79] M. Nardone, V. G. Karpov, D. C. S. Jackson, and I. V. Karpov. A unified model of nucleation switching. *Applied Physics Letters*, 94(10):103509, 2009. p. iv , p. 18 , p. 40 , p. 41 , p. 69

[80] M. Nardone, M. Simon, I. V. Karpov, and V. G. Karpov. Electrical conduction in chalcogenide glasses of phase change memory. *Journal of Applied Physics*, 112(7):071101, 2012. p. 29 , p. 47

[81] R. G. Neale and J. A. Aseltine. The application of amorphous materials to computer memories. *IEEE Transaction on Electron Devices*, 20(2):195–205, 1973. p. 9 , p. 69

[82] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Reviews of Modern Physics*, 81(1):109, 2009. p. 96 , p. 97

[83] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 306(5696):666–669, 2004. p. 96

[84] P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin. Electron transport in disordered graphene. *Physical Review B*, 74(23):235443, 2006. p. 97

[85] S. R. Ovshinsky. Reversible electrical switching phenomena in disordered structures. *Physical Review Letters*, 21:1450–1453, 1968. p. 9 , p. 47 , p. 69

[86] A. E. Owen, J. M. Robertson, and C. Main. The threshold characteristics of chalcogenide-glass memory switches. *Journal of Non-Crystalline Solids*, 32(1):29–52, 1979. p. 47

[87] N. Papandreou, A. Pantazi, A. Sebastian, E. Eleftheriou, M. Breitwisch, C. Lam, and H. Pozidis. Estimation of amorphous fraction in multilevel phase-change memory cells. *Solid-State Electronics*, 54(9):991–996, 2010. p. 109

[88] N. Papandreou, H. Pozidis, A. Pantazi, A. Sebastian, M. Breitwisch, C. Lam, and E. Eleftheriou. Programming algorithms for multilevel phase-change memory. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 329–332, 2011. p. 109

[89] K. E. Petersen and D. Adler. Electronic nature of amorphous threshold switching.

Applied Physics Letters, 27(11):625–627, 1975. p. 42

[90] I. Petrov, R. Imamov, and Pinsker Z. Electronographic determination of the structures of $Ge_2Sb_2Te_5$ and $GeSb_4Te_7$. *Soviet Physics, Crystallography*, 13:339–344, 1968. p. 12

[91] E. Piccinini, A. Cappelli, F. Buscemi, R. Brunetti, D. Ielmini, M. Rudan, and C. Jacoboni. Hot-carrier trap-limited transport in switching chalcogenides. *Journal of Applied Physics*, 112(8):083722, 2012. p. 49 , p. 69 , p. 93

[92] E. Piccinini et al. Novel 3D random-network model for threshold switching of phase-change memories. To appear on *IEDM Technical Digest*, 2013. p. 88

[93] A. Pirovano, A. L. Lacaita, A. Benvenuti, F. Pellizzer, and R. Bez. Electronic switching in phase-change memories. *IEEE Transactions on Electron Devices*, 51(3):452–459, 2004. p. iii , p. 2 , p. 18 , p. 23 , p. 24 , p. 34 , p. 47

[94] E. Pop et al. Sub-10 nm scaling of phase-change memory: Thermoelectric physics, carbon nanotube and graphene electrodes. In *Proceedings of EPCOS*, 2013. p. 96

[95] R. W. Pryor and H. K. Henisch. Mechanism of threshold switching. *Applied Physics Letters*, 18:324, 1971. p. 42

[96] R. W. Pryor and H. K. Henisch. Nature of the on-state in chalcogenide glass threshold switches. *Journal of Non-Crystalline Solids*, 7(2):181–191, 1972. p. 32

[97] P. Ramachandran and G. Varoquaux. Mayavi: 3D Visualization of Scientific Data. *Computing in Science & Engineering*, 13(2):40–51, 2011. p. 70

[98] S. Raoux et al. Phase-change random access memory: A scalable technology. *IBM Journal of Research and Development*, 52(4.5):465–479, 2008. p. 95

[99] S. Raoux, W. Welnic, and D. Ielmini. Phase change materials and their application to nonvolatile memories. *Chemical Reviews*, 110(1):240–267, 2009. p. iii , p. 10 , p. 24

[100] S. Raoux and M. Wuttig. *Phase change materials: science and applications*. Springer, 2009. p. iii , p. 18 , p. 19

[101] A. Redaelli, A. Pirovano, A. Benvenuti, and A. L. Lacaita. Threshold switching and phase transition numerical models for phase change memory simulations. *Journal of Applied Physics*, 103(11):111101, 2008. p. 35

[102] M. Rizzi, M. Ferro, P. Fantini, and D. Ielmini. Energy landscape model of conduction and phase transition in phase change memories. *IEEE Transactions on Electron Devices*, 60(11):3618–3624, 2013. p. 88

[103] M. Rizzi and D. Ielmini. Energy landscape model of conduction and switching in phase change memories. In *IEDM Technical Digest*, pages 589–592. IEEE, 2012. p. 88

[104] D. Roy et al. Contact resistance of TiW to phase change material in the amorphous and crystalline states. In *10th Annual Non-Volatile Memory Technology Symposium (NVMTS)*, pages 12–15, 2009. p. 113

[105] D. Roy, J. H. Klootwijk, D. J Gravesteijn, and R. A. M. Wolters. Contact resistance of tiw to ultra-thin phase change material layers. In *Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, pages 87–90, 2011. p. 113

[106] D. Roy, M. A. A. Zandt, and R. A. M. Wolters. Specific contact resistance of phase change materials to metal electrodes. *IEEE Electron Device Letters*, 31(11):1293–1295, 2010. p. 113

[107] M. Rudan, F. Giovanardi, E. Piccinini, F. Buscemi, R. Brunetti, and C. Jacoboni. Voltage snapback in amorphous-gst memory devices: Transport model and validation. *IEEE Transactions on Electron Devices*, 58(12):4361–4369, 2011. p. 47

[108] D. K. Schroder. *Semiconductor material and device characterization*. John Wiley & Sons, 2006. p. 110 , p. 111 , p. 112 , p. 113

[109] A. Sebastian, A. Pauza, C. Rossel, R. M. Shelby, A. F. Rodríguez, H. Pozidis, and E. Eleftheriou. Resistance switching at the nanometre scale in amorphous carbon. *New Journal of Physics*, 13(1):013020, 2011. p. 69

[110] R. Shanks and C. Davis. A 1024-bit nonvolatile 15ns bipolar read-write memory. In *IEEE International Solid-State Circuits Conference Tech. Digest*, 1978. p. 9

[111] M. P. Shaw, V. V. Mitin, E. Schöll, and H. L. Grubin. *The physics of instabilities in solid state electron devices*. Springer, 1992. p. 31

[112] L. E. Shelimova, O. G. Karpinskii, P. P. Konstantinov, M. A. Kretova, E. S. Avilov, and V. S. Zemskov. Composition and properties of layered compounds in the GeTe-Sb_2Te_3 system. *Inorganic Materials*, 37(4):342–348, 2001. p. 19 , p. 22

[113] A. Shindome, Y. Doioka, N. Beppu, S. Oda, and K. Uchida. Experimental study of two-terminal resistive random access memory realized in mono- and multilayer exfoliated graphene nanoribbons. *Japanese Journal of Applied Physics*, 52:04CN05, 2013. p. 69

[114] B. I. Shklovskii and A. L. Efros. Electronic properties of doped semiconductors. *Moscow Izdatel Nauka*, 1, 1979. p. 48

[115] K. Shportko, S. Kremers, M. Woda, D. Lencer, J. Robertson, and M. Wuttig. Resonant bonding in crystalline phase-change materials. *Nature Materials*, 7(8):653–658, 2008. p. 16

[116] M. Simon, M. Nardone, V. G. Karpov, and I. V. Karpov. Conductive path formation in glasses of phase change memory. *Journal of Applied Physics*, 108(6):064514, 2010. p. 47

[117] G. Spadini, I. V. Karpov, and D. L. Kencke. Future high density memories for computing applications: device behavior and modeling challenges. In *Proceedings of the 15th International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 223–226. IEEE, 2010. p. 1

[118] Z. Sun, J. Zhou, and R. Ahuja. Structure of phase change materials for data storage. *Physical Review Letters*, 96:055507, 2006. p. 13

[119] Z. Sun, J. Zhou, A. Blomqvist, B. Johansson, and R. Ahuja. Formation of large voids in the amorphous phase-change memory $Ge_2Sb_2Te_5$ alloy. *Physical Review Letters*, 102(7):075504, 2009. p. 15

[120] S. Taschini, M. Rudan, and R. Brunetti. Particle and energy fluxes in semiconductors: Full-band hydrodynamic equations and the thermodynamic limit. *Physical Review B*, 60(19):13582, 1999. p. 49

[121] J. Tauc, R. Grigorovici, and A. Vancu. Optical properties and electronic structure of amorphous germanium. *Physica Status Solidi B*, 15(2):627–637, 1966. p. 18

[122] T. Tsafack, E. Piccinini, B.-S. Lee, E. Pop, and M. Rudan. Electronic, optical and thermal properties of the hexagonal and rocksalt-like $Ge_2Sb_2Te_5$ chalcogenide from first-principle calculations. *Journal of Applied Physics*, 110(6):063716, 2011. p. 18

[123] K. Tsendin. Electro-thermal theory of the switching and memory effects in chalcogenide glassy semiconductors. *Physica Status Solidi B*, 246(8):1831–1836, 2009. p. 69

[124] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai. Metal-oxide RRAM. *Proceedings of the IEEE*, 100(6):1951–1970, 2012. p. 69

[125] H.-S. P. Wong, S. Raoux, S. Kim, J. Liang, J. P. Reifenberg, B. Rajendran, M. Asheghi, and K. E. Goodson. Phase change memory. *Proceedings of the IEEE*, 98(12):2201–2227, 2010. p. iv , p. 25 , p. 47 , p. 69

[126] J. D. Wood, S. W. Schmucker, A. S. Lyons, E. Pop, and J. W. Lyding. Effects of polycrystalline Cu substrate on graphene growth by chemical vapor deposition. *Nano Letters*, 11(11):4547–4554, 2011. p. 97

[127] C. D. Wright, Y. Liu, K. I. Kohary, M. M. Aziz, and R. J. Hicken. Arithmetic and biologically-inspired computing using phase-change materials. *Advanced Materials*, 23(30):3408–3413, 2011. p. 3

[128] M. Wuttig and N. Yamada. Phase-change materials for rewritable data storage. *Nature Materials*, 6:824–832, 2007. p. 47 , p. 69

[129] F. Xiong, M.-H. Bae, Y. Dai, A. D. Liao, A. Behnam, E. A. Carrion, S. Hong, D. Ielmini, and E. Pop. Self-aligned nanotube-nanowire phase change memory.

Nano Letters, 13(2):464–469, 2013. p. 4 , p. 26 , p. 31 , p. 70 , p. 93 , p. 95 , p. 107 , p. 108 , p. 110 , p. 112

[130] F. Xiong, A. D. Liao, D. Estrada, and E. Pop. Low-power switching of phase-change materials with carbon nanotube electrodes. *Science*, 332(6029):568–570, 2011. p. v , p. 26 , p. 31 , p. 70 , p. 95 , p. 96 , p. 105 , p. 107

[131] N. Yamada. Erasable phase-change optical materials. *Material Research Society Symposium Bulletin*, 21:48, 1996. p. 47

[132] S. Yamanaka, S. Ogawa, I. Morimoto, and Y. Ueshima. Electronic structures and optical properties of GeTe and $Ge_2Sb_2Te_5$. *Japanese Journal of Applied Physics*, 37(part 1):3327–3333, 1998. p. 23

[133] J. Yao, J. Lin, Y. Dai, G. Ruan, Z. Yan, L. Li, L. Zhong, D. Natelson, and J. M. Tour. Highly transparent nonvolatile resistive memory devices from silicon oxide and graphene. *Nature Communications*, 3:1101, 2012. p. 110

[134] J. Yao, Z. Sun, L. Zhong, D. Natelson, and J. M. Tour. Resistive switches and memories from silicon oxide. *Nano Letters*, 10(10):4105–4110, 2010. p. 69 , p. 110

[135] J. Yao, L. Zhong, D. Natelson, and J. M. Tour. In situ imaging of the conducting filament in a silicon oxide resistive switch. *Scientific Reports*, 2, 2012. p. 110

[136] V. V. Zhirnov, R. K. Cavin III, J. A. Hutchby, and G. I. Bourianoff. Limits to binary logic switch scaling-a gedanken model. *Proceedings of the IEEE*, 91(11):1934–1939, 2003. p. 47