A Dissertation

entitled

Second Phase Filamentation and Bulk Conduction in Amorphous Thin Films

by

Mark A. Simon

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Physics

Dr. Victor G. Karpov, Committee Chair

Dr. Robert Deck, Committee Member

Dr. David Ellis, Committee Member

Dr. Sanjay Khare, Committee Member

Dr. Daniel Georgiev, Committee Member

Dr. Patricia R. Komuniecki, Dean
College of Graduate Studies

The University of Toledo
May 2011

An Abstract of

Second Phase Filamentation and Bulk Conduction
in Amorphous Thin Films

by

Mark A. Simon

Submitted to the Graduate Faculty as partial fulfillment of the requirements
for the Doctor of Philosophy Degree in Physics

The University of Toledo
May 2011

In this work the phenomenon of second phase conductive filaments is explored, with a focus on device applications. The second phase is often characterized by a significant increase in electrical conductivity, e.g. the result of the transition from an amorphous to a crystalline state in chalcogenide materials or VO_2 used in modern electronic devices. A filament is formed when a region consisting solely of the second phase spans the volume of a host material. Such filaments may either be persistent or reversible depending on material parameters and operating regime underlying their formation.

When two opposing sides of the host material have electrical contacts, a conductive filament between the two will act as shunt. The presence of a shunt drastically reduces the device resistance. In memory applications a shunt that forms unintentionally will lead to data loss, while at other times it is necessary to form a shunt in order to produce a detectable change in the state of the system.

Theories are developed to describe the processes behind both types of filamentation. For each an underlying theoretical framework is presented, and then analytical methods are used to derive key results in terms of material parameters and device geometry. In doing so there is a focus on thin-film devices with inter-electrode distances that can be as small as tens of nanometers. Numerical simulations are also

employed to substantiate the analytical results. Lastly, a comparison is made showing agreement with the available experimental data.

The theories presented here are rather general in nature. In order to provide examples of practical applications and to compare with experimental data, the discussion is usually developed within the context of chalcogenide glass switches. Such switches have found applications as memory devices, of which there are two distinct types: phase change memory (PCM) and threshold switches (TS).

Along with the work on second phase formation discussed above, a survey of conduction mechanisms in bulk chalcogenide glasses is also presented. In addition to the existing models of conduction, our consideration here proposes two new mechanisms: percolation conduction in a potential relief created by second phase particles, and pinhole conduction channels through very thin structures. Many of the results are equally applicable to other disordered systems. The motivation behind the review is to summarize the established physics of charge transport in bulk chalcogenide systems that can exist in parallel with the second phase filament or can precede the filament formation in chalcogenide glasses. We point out potential shortcomings in our current level of understanding and suggest specific relationships that may be obtained from future experiments which can be used to indicate which mechanism(s) are dominant.

Dedicated to my good friend Felix

Acknowledgments

I would like to thank the University of Toledo Department of Physics and Astronomy, the previous department chair Dr. Karen Bjorkman and current chair Dr. Lawrence Anderson, and all of the faculty for their exceptional instruction over the years. I especially thank the members of my committee for their support and guidance. I would also like to thank the office staff, whom are always so eager to help. In particular Richard Irving, who has assisted me on a number of occasions.

I've wanted to be a physicist for as long as I can remember, even as a young child. I thank my parents, Lisa and Randy for always being supportive in my endeavors and also my sisters Alisia and Elizabeth for all the good times we've had together.

Many thanks to Dr. ILya Karpov for directing financial support through the Intel Corporation and for his hospitality during my visit to Intel's headquarters. A special thanks to Marco Nardone for all of his hard work put into our collaborations. Best of luck to you in your future pursuits. Most of all, I would like to thank my advisor Dr. Victor Karpov. Over the course of our research he has secured funding which has allowed me to present at Materials Research Society meetings. Much more importantly, he has imparted a great deal of knowledge derived from his years of experience and has always encouraged me to perform at my fullest. I look forward to many years of future collaboration as my career progresses.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Filament formation plays in important role in many modern device applications. Shunts consisting of a highly conductive phase may result in device degradation, manifested as decreased performance or data loss. The probability of shunt formation is usually derived within the framework of percolation theory. Once the volume fraction of conductive material reaches the percolation threshold, a cluster is formed which is able to span the device infinitely in all directions, regardless of device dimensions. In Chapter 2 we discuss the limitations of percolation theory when applied to sufficiently thin devices. Here it is shown that the concept of a percolation cluster need not apply: relatively short chains of conducting particles are able to shunt the device well below the percolation threshold. Analytical methods are used to describe the shunting probability in terms of key material parameters and device dimensions and a numerical model is presented which confirms our results.

The model developed for shunt formation is then applied to the specific case of phase change memory (PCM). Different modes of device failure are considered such as crystal growth and the nucleation of conducting particles within the amorphous host. The time and temperature dependence of each is derived as well as the corresponding activation energy.

In Chapter 3 we model the thermodynamics of a conducting filament under an external bias. A general framework is introduced which describes the growth of a

filament in terms of the free energy of the system. A simple model using the geometry of a flat plate capacitor is described in order to provide closed-form analytical results. Numerical simulation software using the finite element method is then used to model a specific threshold switch (TS) device. Excellent agreement with the available experimental data is achieved.

A survey of conduction mechanisms in chalcogenide glasses is presented in Chapter 4. Although written within the context of chalcogenide systems, many of the results may also be applied to similar disordered materials. First, a brief overview of the experimental data is provided which describes the commonly observed nonlinear current-voltage characteristics. Next, a number of possible mechanisms for charge transport are discussed including:

1. Poole-Frenkel effect
2. Schottky decrease in interfacial barrier near device electrodes
3. Field-induced delocalization of shallow band tail states near the mobility edges
4. Space charge limited (injection) currents
5. Field effects in hopping conduction
6. Percolation conduction
7. Conduction through crystalline inclusions in an amorphous matrix

With each mechanism analytical expressions for the conductivity are given as well as the the range of applied field under which their contribution may be significant.

Although quite general in nature, much of the theory presented here is applied to chalcogenide glass switches. This provides a specific real-world application of the theory and allows for a comparison with experimental data. The following is a brief overview of PCM and TS operation which includes a description of how our models of filamentation in each can be used to solve specific problems.

1.1 Overview of Chalcogenide Based Memory

Devices containing chalcogenide material have long been used for memory applications due to their ability to reversibly switch between two distinct phases, both quickly and reliably. The amorphous phase is characterized as having a significantly lower conductivity and reflectivity compared to the crystalline phase. Binary information is encoded by initiating a phase transformation between the two states and the information is then read by detecting which phase the sample is in. The most familiar example is optical recording media, where a laser is used to induce a phase change across a small area of a disc composed of a chalcogenide alloy. The same laser, operating at a lower power, is then used to detect the difference in reflectivity.

Similar to optical recording media, both types of chalcogenide glass switches also make use of the transition between the amorphous and crystalline phase to encode information. Since the resistivity of the two phases differs by orders of magnitude, by applying a small bias to the sample one can determine the device resistance and hence which state the sample is in. In order to initiate the amorphous to crystalline transition, the applied bias needs to exceed a critical value referred to as the threshold voltage (V_{th}). Once the threshold is reached, the device will quickly transition into the crystalline state.

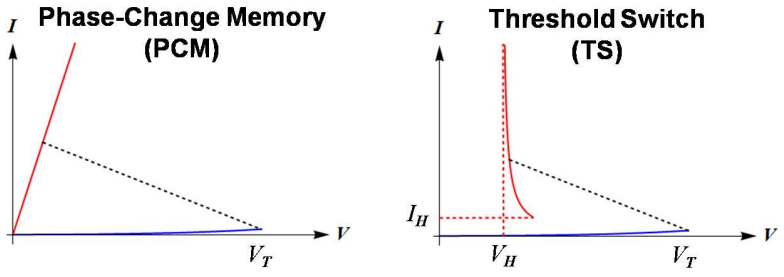

Figure 1-1: Qualitative current voltage characteristics of PCM and TS.

TS require a minimum current, referred to as the holding current (I_h), in order

to sustain the crystalline phase. If $I < I_h$, the device will quickly switch back to the amorphous phase. In the case of PCM, the crystalline state remains after field removal and a relatively larger current needs to be applied in order to reset the device back into the amorphous state.

There are a number of models for the switching process that can be categorized as being either thermal or electric in nature. Since the mid-1970s the consensus has been in favor of electronic models to describe switching in thin films. Recently, a theory of field-induced nucleation has been proposed which is able to explain a number of experimental observations, such as the similarity in switching data for PCM and TS.

1.1.1 Data Retention Problem in Phase Change Memory

When a PCM device is read, a small voltage is applied to the sample in order to determine its resistance. The amorphous phase is metastable and will crystallize over time, especially at elevated temperature. If a chain of crystalline particles connects the two electrodes, a shunt will form and the device resistance will decrease significantly. When the cell is read, a shunted device initially in the amorphous phase will appear as if it was set to the crystalline state. The crystalline phase, on the other hand, is stable and a device in the crystalline state will not be misread unless it was improperly written. Therefore, by finding the probability of shunt formation one can determine the likelihood that a device will fail.

1.1.2 The ON State of Threshold Switches

Once the voltage across a TS exceeds the threshold value, the device transitions into the ON state. The device will remain in the ON state as long as the current is maintained above a critical level. There is evidence to suggest a highly conductive channel (or filament) exists within the device when it is in the ON state. A number of experimental techniques have been used to characterize the radius r of the filament

as a function of current. The relationship $r \propto \sqrt{I}$ is often cited, but lacks a sound theoretical basis. We propose a thermodynamic model which provides this functional dependence in terms of key material parameters. The model can also be applied to explain other operating characteristics such as the nature of the holding current and can be extended to model transient behavior.

1.1.3 Bulk Conduction in Chalcogenide Glasses

The effects of second phase filaments can be adequately recognized when separated out from the background bulk conduction of the amorphous matrix. While general principles of charge transport in disordered systems apply fully to the case of chalcogenides, those systems have specific features distinguishing them into a separate class of materials. In particular, it was found that hopping conduction is significantly suppressed (practically nonexistent under room temperature) and non-ohmic effects are very strong and exhibit themselves clearly at the stages preceding the phase transformations.

In general, the understanding of the unique transport properties of chalcogenide glasses is based on Anderson's postulate of anomalously strong electron-lattice interactions and related negative correlation energy leading to the pairing of localized electrons. It has been established that low-field conduction is due to transport at mobility edges. However the mechanisms of non-ohmic conduction under moderately strong electric fields $\gtrsim 10$ kV/cm remained poorly understood, and the researchers were split between the classical Poole-Frenkel interpretation and a recently revived hypothesis of hopping. This uncertainty was aggravated by a strong competition between different groups trying to establish priority in the newly unfolding area of technological significance - phase change memory. As a result of that competition, a number of newcomers in this area started to reinvent the physics of glasses putting forward hypotheses totally disconnected from the earlier achieved understanding. This

led to even stronger confusion slowing down further progress in PCM research. It is our approach here that clarity can be attained by carefully reviewing all the conceivable mechanisms of non-ohmic transport and proposing which of them can be adequate and how further clarification can be advanced. Chapter 4 of this thesis is devoted to this task.

Chapter 2

Conductive Path Formation in Glasses of Phase Change Memory

Chalcogenide phase change memory (PCM) holds promise as the next generation of non-volatile memory.[1, 2] Memory is enabled by two markedly different resistance states that can be repeatedly achieved by appropriate voltage pulses; they are a highly resistive amorphous phase and a low resistance crystalline phase. The amorphous phase is metastable and tends to crystallize over time at elevated temperatures, which causes either shunting of a PCM cell or a significant drop of amorphous phase resistance, resulting in data loss. Data retention characteristics quantify the ability of a device to maintain data for extended periods of time at elevated temperatures. Estimation of the failure time is typically performed by measuring the time dependent resistance change under constant temperature bake. While the concept of data retention related to crystallization is generally accepted, the physics and underlying mechanisms are still poorly understood and insufficiently modeled. In the present work, we develop a theoretical model that describes the probability and statistics of device failure in terms of device parameters and external conditions. Our model goes beyond the traditional percolation theory approach and shows the limitations of the latter with respect to thin-film technologies.

In addition to PCM, there exist other structures in modern technologies where

conductive path formation is an important mode of device degradation. Thin oxide films in metal oxide semiconductor field-effect transistors (MOSFETs) represent another such example where shunting plays a significant role in reliability. During the course of operation these devices accumulate point defects which can form a shunting path between the gate and semiconductor material. [3, 4] A third example relates to photovoltaic devices where shunting remains a major degradation mode. [5, 6]

The theoretical approach to shunting usually employs the percolation concept, [4, 7, 8] modeling the device as a lattice with spacing $2R$ where R is the effective radius of a conductive particle [Fig. 2-1(a)]. The conductive particles randomly occupy a fraction v of the lattice sites and form clusters, some of which can create a connected path that shunts through the film. For PCM, a typical value is $R \sim 3$ nm corresponding to the nucleation radius for the crystalline phase. For thin oxides, there is some debate as to the defect size, [3] but it is generally agreed upon to lie between $0.5 - 1.5$ nm. The practically interesting film thicknesses are in the range $1 - 100$ nm.

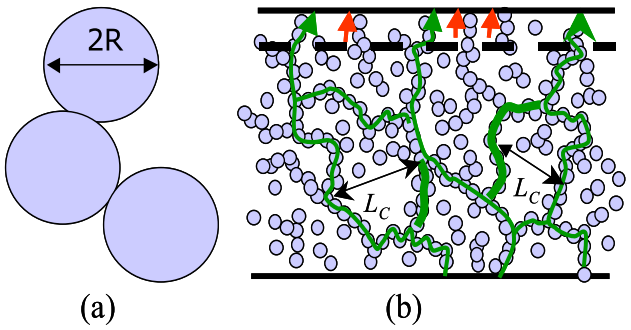

Figure 2-1: (a) A fragment of the shunting path formed by crystalline nuclei of radius R; and (b) the topology of the infinite percolation cluster (b) with a backbone shown in green. The thick dashed line represents the case of a very thin structure. L_c shows the correlation radius, which can be identified with the characteristic mesh size of the infinite percolation cluster with $v > v_c$.

The concept of a percolation cluster implies a connected structure that is infinite in space and capable of establishing conduction between any remote points; such a

structure exists when the volume fraction occupied by the conductive phase exceeds a certain critical value[9] $v_{c\infty} \sim 0.3$. The percolation cluster is characterized by its correlation radius L_c (Fig. 2-1). Near the percolation threshold,[9] $L_c \sim 2R|v-v_{c\infty}|^{-\nu}$ with $\nu \approx 0.9$. For $v > v_{c\infty}$, the correlation radius can be geometrically interpreted as the characteristic mesh size of the cluster, so the latter behaves as a macroscopically uniform medium when the systems size is much greater than L_c, while it has a complex topology on scales less than L_c. For $v < v_{c\infty}$, L_c would describe the characteristic gyration radius of finite clusters that do not form connected infinite pathways; that latter case is irrelevant to the purpose of this work.

For thin films, an infinite percolation cluster is not necessary to form a shunt when its correlation radius will exceed the film thickness L (for v close to $v_{c\infty}$). Therefore, in the domain of $L_c \gg L$, the percolation cluster backbone will play almost no role against the background of multiple finite clusters connecting the electrodes as illustrated in Fig. 2-1(b). The concept of percolation theory for thin films becomes applicable when

$$v > v_{c\infty} + (2R/L)^{1/\nu} \equiv v_c. \tag{2.1}$$

Geometrically, the latter inequality means that the volume fraction occupied by the crystalline phase is high enough to make the cluster mesh size smaller than the film thickness. We note that the size dependent term in Eq. (2.1) can be significant for modern thin-film devices. For example, taking $R = 3$ nm and $L = 30$ nm results in $2R/L = 0.2$; comparable to $v_{c\infty}$. In this example, the percolation cluster scenario of shunting ceases to work already at $v \lesssim 0.5$, which is in the entire range of practical interest.

Here, we concentrate on the range of $v < v_c$ which is important because it determines the statistics of early failures (before the percolation cluster is formed).[10, 11] We show that linear chains nearly perpendicular to the electrodes [short arrows in Fig. 2-1(b)], which are not part of the percolation cluster, are responsible for shunting in

that regime. We denote such chains as rectilinear pathways (shunts). Prior work [4, 3] described the pre-threshold range by means of computer simulations. Our analytical approach here (briefly introduced in Ref. [12]) is more explicit by explaining the nature of the deviations from percolation theory and predicting early failure statistics. We note that the first work that recognized the failure of the percolation approach [13] for thin films was related to the problem of transversal hopping conduction in amorphous films; a review of subsequent work on the topic presented in Ref. [14] was inspirational for the present project.

Our paper is organized as follows. We start with a numerical experiment in Sec. 2.1 that naturally introduces the motivation and the concept of this work. In Sec. 2.2 we present a general analytical approach to the probabilistic description of early fail shunting. Section 2.3 specifies our approach to the cases of different device geometries. These results are translated into temporal dependencies of shunting probability in Sec. 2.4. Possible effects of external fields on the conductive path and early fail probability is described in Sec. 2.5. Finally, Sec. 2.6 presents the conclusions of this work.

2.1 Numerical Experiment

As a motivation for the analytical effort that follows, we begin with a numerical experiment involving 3D percolation on a simple cubic lattice. We considered a system of 10 by 10 by L lattice points where L is the device thickness in lattice units $2R$ (corresponding to the particle diameter). The sites of the lattice were randomly populated with conducting particles up until a given volume fraction v. When the nearest-neighbor of an occupied site is also occupied, they are considered to be connected and part of the same cluster. The condition for shunting is that there exists a cluster which spans the device (i.e., shunting occurs when a cluster starts at the base of the system and ends at the opposing side a distance L away).

For each parameter setting, 1000 trials were run. For each trial, it was determined

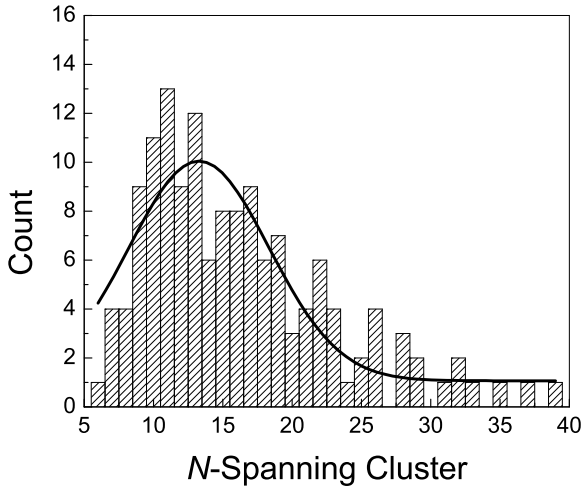

Figure 2-2: Numerically modeled statistics of the number of nuclei N in shunting paths for a 10x10x5 cube with $v = 0.2$. Spanning cluster formation occurred 143 times in 1000 trials. The curve represents the Poisson distribution.

whether a shunt formed and, if so, the number N of particles in the spanning cluster. Fig. 2-2 reveals a maximum in the probabilistic distribution of N observed for not too small $v > 0.1$. As shown in Fig. 2-3(b), the most probable number of particles N_0 in a shunting pathway was found to be linear in the film thickness (L), implying that these pathways are almost linear and perpendicular to the electrodes (for comparison, random walk pathways would have $N_0 \propto L^2$). For extremely small $v < 0.1$, we mostly observed exactly linear pathways of length L. Also, our modeling results in Fig. 2-3(a) show the linear area dependence of the shunting probability.

The above findings can be attributed to shunting by nearly rectilinear chains of particles [Fig. 2-4(b)] with a relatively insignificant degree of winding compared to the percolation cluster topology. The maximum in Fig. 2-2 can be understood as arising from two competing tendencies: 1) larger clusters are more likely to be able to form a shunt and 2) the probability of finding an N-particle cluster decreases with cluster size N.

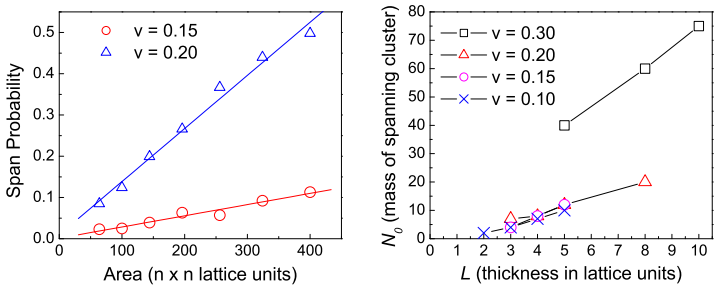

Figure 2-3: (a) Modeling results for the area dependence of the shunting probability. (b) Numerical simulation of the typical number of sites in the shunt (N_0) in 10x10xL unit systems for various volume fractions v. Each data point represents the average of 1000 trials. Note that for $v = 0.3$ there is a strong increase in N_0 which is indicative of non-rectilinear paths and the percolation cluster scenario.

Next we give an analytical description of such nearly rectilinear chains and their statistics, in agreement with the results of the numerical experiment. Note that from this point on, all figures (other than plots) are qualitative sketches and should not be directly compared with the numerical modeling discussed above.

2.2 Analytical Model of Shunting Paths

2.2.1 General Approach

For a structure with bottom electrode area $A \gg R^2$, the probability of forming a chain of N particles shunting through thickness L is given by $p(N, L) = p_N(L)p_v(N)$. Here $p_N(L)$ is the conditional probability of an N-particle chain to have projection of length L onto axis z perpendicular to the electrode; $p_v(N)$ is the probability to form an N-particle chain in a system where randomly distributed particles occupy volume fraction v. The total probability of shunting through the film is

$$P(L) = \frac{A}{(2R)^2} \int_{L/2R}^{\infty} p_N(L)p_v(N)dN, \tag{2.2}$$

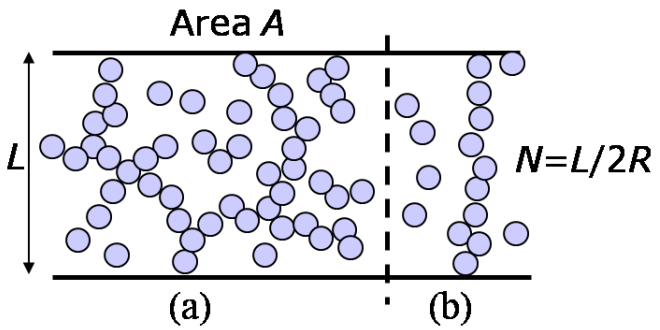

Figure 2-4: Two scenarios of shunting between two electrodes: (a) percolation; and (b) through a nearly rectilinear path characterized by a random walk with the displacement roughly equal to the distance L between the electrodes.

where the lower limit of integration is the minimum number of particles required for shunting. The multiplier $A/(2R)^2$ gives the number of possible shunting pathways starting at one electrode, providing an explanation for the linear dependence of the shunting probability on the area, as observed in the numerical experiment [Fig. 2-3(a)]. If $P(L)$ in Eq. (2.2) is larger than 1 it is to be interpreted as the number of shunting pathways formed.

Assuming the nucleation events to be spatially uncorrelated, the probability of forming a chain of N particles long may be represented by the Poisson distribution

$$p_v = \overline{N}^N e^{-\overline{N}}/N!$$

where $\overline{N} = Nv$ and $v \ll 1$ is the volume fraction of the crystalline phase. Using Stirling's approximation, this may be rewritten in the form [15]

$$p_v = \frac{e^{-N\Gamma}}{\sqrt{2\pi N}}, \quad \Gamma \equiv \ln\left(\frac{1}{v}\right) + v - 1. \tag{2.3}$$

Chains of randomly dispersed conductive particles have the topology of a random walk (Fig. 2-4) with step size $2R$ in a half-space[16] starting at one electrode . The

normalized probability of the chain ending at the other electrode is

$$p_N = \frac{3L}{2NR} \exp\left(-\frac{3L^2}{8NR^2}\right), \tag{2.4}$$

where R is the effective nucleus radius and L the distance between electrodes (for a derivation, see Appendix A). We note that, strictly speaking, the model of self-avoiding random walk would be a more adequate description of nuclei formed chains. It will however become apparent later that the degree of chain winding of practical interest is relatively insignificant (chains are close to rectilinear). The probability of self-crossing is small enough so that it may be safely neglected in favor of a simpler mathematical analysis.

In the composite probability $p = p_N p_v$, the two multipliers exhibit opposite trends favoring larger and smaller N's respectively. As a result, the composite probability obtains a maximum at a particular number of particles in the chain, denoted by

$$N_0 = \frac{L}{2R}\sqrt{\frac{3}{2\Gamma}}, \tag{2.5}$$

obtained by a straightforward optimization of the composite exponent. This provides an explanation for the maximum in the probabilistic distribution of Fig. 2-2.

N_0 must always be greater than the minimum number of particles spanning the device, $L/2R$. For lower volume fractions, one has to use $N_0 = L/2R$ (rectilinear paths). By setting $N_0 = L/2R$ in Eq. (2.5), we find the volume fraction $v_0 \sim 0.09$ for which the most likely number of particles in the shunt is the minimum physically allowed. For volume fractions in the range $v_0 < v < v_c$, the most probable number of particles in the connecting chain is between $L/2R < N_0 < 1.5L/2R$. That is, shunting is due to nearly rectilinear paths which depend linearly on thickness, in agreement with the results of the above numerical modeling [Fig. 2-3(b)].

Depending on where the peak of the composite probability lies with respect to

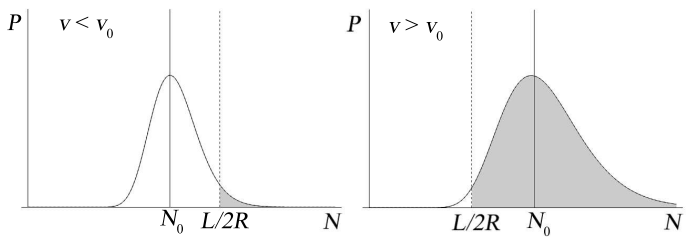

Figure 2-5: The composite probability (P) for shunting as a function of the number of particles in a chain (N). Depending on the volume fraction being considered, the calculated most likely number of particles given by Eq. (2.5) may lie below the physical limit $N = L/2R$.

the lower limit of integration, different approximations to the integral in Eq. (2.2) may be employed (refer to Fig. 2-5). For $v < v_0$ the integral is best approximated by the integrand evaluated at $N = L/2R$, while for $v_0 < v < v_c$ we use the method of steepest descent yielding,

$$P = \begin{cases} \dfrac{3A}{4R^2}\sqrt{\dfrac{R}{\pi L}}\exp\left[-\dfrac{L}{2R}\left(\dfrac{3}{2}+\Gamma\right)\right] & \text{for } v < v_0 \\ \dfrac{A}{R^2}\sqrt{\dfrac{3}{64}}\exp\left(-\dfrac{L}{R}\sqrt{\dfrac{3\Gamma}{2}}\right) & \text{for } v_0 < v < v_c. \end{cases} \tag{2.6}$$

2.2.2 Critical Thickness and Critical Area

Setting $P = 1$ gives the volume fraction that will cause the device to shunt with certainty for a particular thickness. Setting then $v = v_c$ in the equation $P = 1$ we define the critical thickness

$$L_c = R\sqrt{\frac{2}{3\Gamma(v_c)}}\ln\left(\frac{A}{R^2}\sqrt{\frac{3}{64}}\right), \tag{2.7}$$

such that devices of thickness $L < L_c$ are typically shunted by rectilinear paths, while the percolation cluster scenario dominates for $L > L_c$. Similarly, a critical area can be defined such that for a given thickness, systems with $A > A_c$ are typically shunted by rectilinear paths, while the percolation cluster scenario dominates for $A < A_c$,

where

$$A_c = \frac{8R^2}{\sqrt{3}} \exp\left(\frac{L}{R}\sqrt{\frac{3\Gamma(v_c)}{2}}\right). \tag{2.8}$$

As a numerical example, for a PCM cell with $R \sim 3$ nm and $A \sim 2500$ nm^2, we have $L_c \sim 14$ nm. By setting $L \sim 45$ nm (within the range of current PCM thicknesses), we find $A_c \sim 19$ μm^2.

2.2.3 Small Area Cells

It has been tacitly implied in the above that the cell area is large enough to rule out any edge effects. Here we discuss the conditions under which the edge effects can be either accounted for or neglected.

If the side surface of the cell is chemically or structurally different from the bulk, then the barrier for nucleation along the side surface can be lower than that of the bulk. Nucleation will then preferentially evolve in the side interfacial layer with characteristic cross-sectional area $A = Rl$ where l is the perimeter of the cell, which should be used instead of the cell area in the above results. This modification will not lead to any new qualitative features.

In the opposite case of the interfacial nucleation barriers greater than that of the bulk, the side surface will play the role of a repulsive wall for the random walk trajectories dealt with in the above. While that repulsion could change the random walk topology, its effect is extremely small in cells of practically interesting dimensions as shown next.

Given the optimum number of particles in Eq. (2.5) the corresponding number of random walk steps parallel to the electrode plane is estimated as

$$\delta N_0 = \frac{L}{2R}\left(\sqrt{\frac{3}{2\Gamma}} - 1\right). \tag{2.9}$$

The cross-sectional area of such random walk $\delta A = (2R)^2 \delta N_0$ must be much smaller than the cell area A in order to neglect the side surface effects, i. e.

$$\frac{2R}{L}\left(\sqrt{\frac{3}{2\Gamma}} - 1\right) \ll \frac{A}{L^2}. \tag{2.10}$$

Based on the above consideration [see the discussion after Eq. (2.5)], the term in parenthesis is smaller than 0.5 for volume fractions below the percolation threshold. Taking into account that $R \ll L$, the inequality in Eq. (2.10) is obeyed even for cells with an effective area noticeably smaller than thickness. For practical devices with relatively higher aspect ratios ($A \gtrsim L^2$), the inequality in Eq. (2.10) is more strictly adhered to. The physical meaning of such a mild condition on the cell area is that the paths under consideration are close to rectilinear and exhibit relatively insignificant winding; hence, a suppressed interaction with side surfaces. The side surface effects could be more pronounced for $v \sim v_c$, which is beyond the scope of this work.

2.2.4 Extension to Large Arrays

The probability for an individual cell failure described by Eq. (2.6) remains small for most practical device dimensions. However, for technology applications it is important to determine the statistics of large arrays of $n \gg 1$ PCM cells. Let P_0 describe the probability of an individual cell to fail. Assuming all cells in the array have the same probability of failure (which is true if they have roughly the same dimensions and volume fraction of conducting particles), then the probability that one cell in the array will fail is given by $P = 1 - (1 - P_0)^n \approx nP_0$.

Intuitively, the probability of failure for one cell in an array may be obtained directly by replacing the factor A in Eq. (2.6) with the product An. That is, the failure probability is given by the total number of potentially shunting sites in the array, which in turn is proportional to the total effective area (regardless of device geometry). For example, by replacing A with An in Eq. (2.7) and using the same values as before,

the thickness required for failure of one cell in an array of $n = 10^8$ cells at volume fractions below the percolation threshold is $L_c \sim 75$ nm, which is above modern device thicknesses. Hence, rectilinear pathways, rather than percolation clusters, will be the dominant shunting mechanism for thinner devices.

2.3 Device Geometry

As discussed in Sec. 2.2, we model the formation of shunting pathways as a 3D random walk. For simplicity, the analytical model presented throughout the paper thus far has been that of a parallel plate configuration with the amorphous region extending infinitely in the lateral directions [Fig. 2-6 (a)]. In this section we will discuss the effect that a non-flat geometry has on the shunting probability.

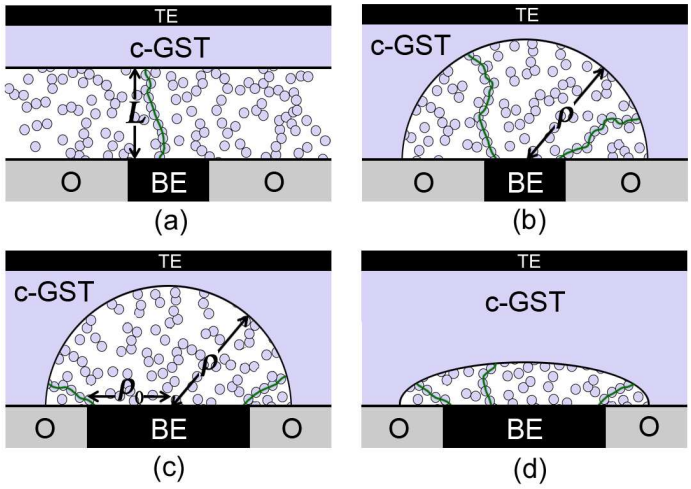

Figure 2-6: Different geometries for the amorphous region lying above a bottom electrode (BE) and oxide (O) layer. Due to its relatively high conductivity, the surrounding crystalline chalcogenide (c-GST) essentially acts as a top electrical contact (TE). (a) Parallel-plate configuration, (b) hemi-sphere with relatively small electrode area, (c) hemi-sphere with electrode area comparable to the base area of amorphous region, and (d) half-ellipsoid. In each, the most likely conducive pathways are given by green lines.

After the initial fabrication, a typical PCM device will consist of a crystalline

region of chalcogenide material sandwiched between a top electrical contact and a bottom layer which contains an electrode surrounded by an insulator (known as the "lance configuration" [17]). A programming current is then applied which raises the temperature of the chalcogenide material through joule heating. An amorphous volume forms near the bottom electrode due to the relatively higher current density in that region, the 'dome' ('mushroom') shape of which [18] depends on applied voltage, electrode area, and crystalline layer thickness.

As our first case, consider a hemi-spherical amorphous volume with radius ρ. If the area of the bottom electrode is small compared to the area of the amorphous region [Fig. 2-6 (b)], we may assume that the shunting probability is the same regardless of where on the electrode the path begins. Instead of Eq. (2.4), the probability for an N-step random walk to end at the conductive crystalline layer is given by,

$$P_N = \frac{9\rho^3}{16N^2R^3} e^{-3\rho^2/8NR^2}. \tag{2.11}$$

This is essentially the same as Eq. (2.4), differing only in the pre-exponential and with the radius of the amorphous region taking the place of the thickness L used in the parallel plate case.

Next we consider an amorphous volume similar to the previous, but with an electrode of area comparable to that of the amorphous region [Fig. 2-6(c)]. The main contribution to shunting will come from a ring of sources located near the edge of the electrode resulting in the N-step probability,

$$p_N = \frac{\rho}{\rho_0}\sqrt{\frac{3}{2\pi N}} e^{-3(\rho-\rho_0)^2/8NR^2}, \tag{2.12}$$

with the approximation that $\rho_0\rho \gg 4R^2$, where ρ_0 is the radius of the electrode. This is once again, similar to Eq. (2.4), but with the effective thickness being the smallest distance between the bottom electrode and surrounding crystalline region $\rho - \rho_0$.

We note that a device initially with a small electrode area (compared to the amorphous region) will eventually become the present case due to growth of the surrounding crystalline material (see Sec. 2.4.1). In other words, over time, the case of Fig. 2-6(b) becomes that of Fig. 2-6(c).

Unlike Eqs. (2.4) and (2.11), Eq. (2.12) cannot be directly placed into Eq. (2.2) to determine the total shunting probability because the number of sites contributing is no longer from the entire area $A/4R^2$, but rather only from the circumference $\pi\rho_0/R$. It is sufficient to consider only the sites located at the edge; sites away from the edge will contribute exponentially less as can be seen from Eq. (2.12).

Lastly, we mention the case of when the surrounding crystalline layer is a constant distance from the bottom electrical contact [Fig. 2-6 (d)]. The result is the same as Eq. (2.12) above (other than the pre-exponential); the main difference is that now the entire electrode area can be considered when calculating the total shunting probability, changing the result by the multiplier ρ_0/R.

We would like to emphasize that the device geometry affects the shunting probability by changing what is considered to be the relevant thickness. Shunts are more likely to occur from sites on the electrode which minimize the distance between the bottom electrode and the surrounding crystalline region. Within the intended accuracy of our model, this minimum distance can replace the thickness L used throughout the paper for the parallel plate case.

2.4 Time and Temperature Dependence of Shunting Probability

Eq. (2.6) describes the time and temperature dependence of the shunting probability through the parameters L, R, and v. There are three cases to consider: 1) the decrease of the amorphous region thickness L due to the growth of the surrounding

crystalline layer, 2) the radial growth (increase in R) of existing conducting particles in the system, and 3) the addition of conducting particles to the system through nucleation (increase in v); each will be treated separately. We conclude this section by calculating the activation energies for each of the above failure mechanisms and comparing to experimental data.

2.4.1 Decay of the Amorphous Region

The amorphous region of a PCM device is surrounded by a crystalline layer which grows over time (see Fig. 2-7). This can be due to either crystal growth or preferential nucleation along boundary surfaces (such as the substrate/amorphous boundary or the crystalline/amorphous boundary). Ignoring the addition and growth of crystalline particles away from the boundaries, the volume fraction will remain constant in time.

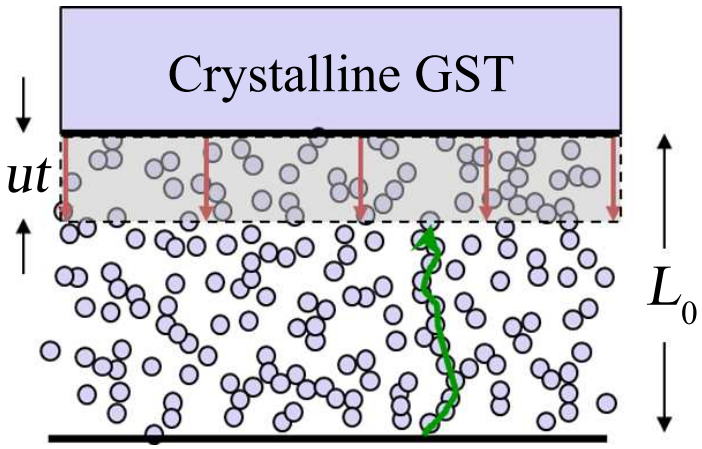

Figure 2-7: The effect of crystalline layer growth on the possible conduction pathways. A chain formerly unable to shunt the device can eventually become a shunt as the thickness decreases (green line).

The initial volume fraction v_i may be the result of the melting-to-freezing transition of the reset operation. As the chalcogenide cools after a reset pulse, there may be enough time to allow for the organization of small crystalline phase regions within the material.

All that is necessary to take into account the growth of the crystalline layer is to transform L in Eq. (2.6) to a function of time and temperature. If we assume that growth is linear with time so that $L = L_0 - ut$ (where L_0 is the initial amorphous thickness and u is the linear growth rate), form Eq. (2.6) we have

$$P \propto \exp\left[\frac{ut}{2R}\left(\frac{3}{2} + \Gamma(v_i)\right)\right]. \tag{2.13}$$

It should be noted that there is an upper limit to the time that can be used in the above equation. The total growth must be less than the initial thickness ($ut < L_0$).

2.4.2 Radial Growth of Crystalline Particles

Next we consider the case of a pre-existing volume fraction v_i of crystalline particles which grow radially over time with no additional nucleation taking place (Fig. 2-8). We again assume that the growth rate is linear so that the radius of a conducting particle is given by $R = R_0 + ut$. The volume fraction is given by $v = v_i(R/R_0)^3$ and hence the total shunting probability is

$$P \propto \exp\left[-\frac{L}{ut}\sqrt{\frac{3}{2}\left(-\ln v_i - 3\ln ut/R_0 - 1\right)}\right]. \tag{2.14}$$

As before, there is a limitation to the timeframe allowed. Certainly, the total radial growth should be less than the device thickness ($ut < L$), but it should also be low enough so that the volume fraction stays below the percolation threshold v_c. This may be expressed as $ut < R_0(v_c/v_i)^{1/3}$.

Presented in Fig. 2-9 are the failure statistics plotted in the Weibull coordinates commonly used in reliability studies. The radial growth was limited so that the volume fraction remained below the percolation threshold. As the threshold is reached, the failure probability approaches the top x-axis (corresponding to 90% failure), indicating that the cell will fail with near certainty. Notice that increasing the thickness

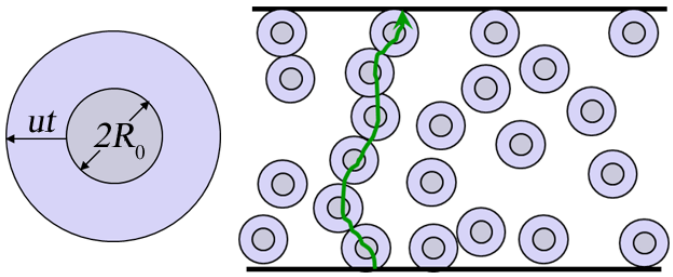

Figure 2-8: The effect of the radial growth of a pre-existing crystalline volume fraction. A chain formerly unable to shunt the device is able to once the particles reach a certain size (green line).

drastically changes the slope and intercept, implying that thinner devices have much shorter retention times. Reducing the temperature shifts the curves to the right, indicating the longer retention times at lower temperatures.

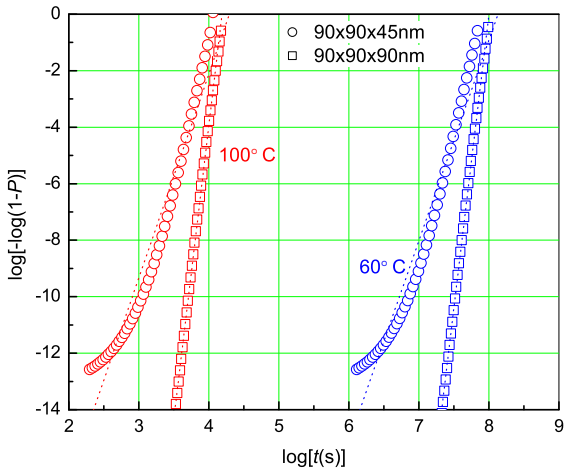

Figure 2-9: Weibull plots of failure statistics for an initial volume fraction $v_i = 0.01$ of crystalline particles growing radially over time in accordance with Eq. 2.14. Data from Ref. [19] was used for the growth rate u.

2.4.3 Nucleation Driven Models

Crystalline particles will nucleate within the amorphous region resulting in a time dependent volume fraction $v(t)$. Here we discuss two possible models for the time and temperature dependence.

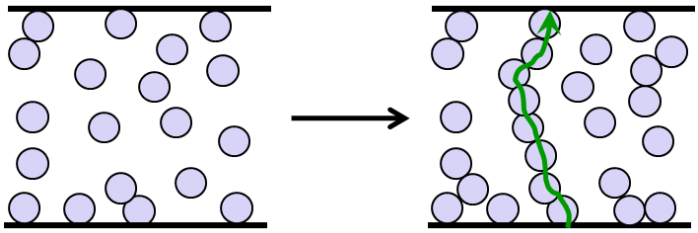

Figure 2-10: The nucleation of crystalline particles will increase the volume fraction and create shunting pathways (green line).

Our first model deals with a fixed nucleation barrier that does not vary between different microscopic regions in a glass. The steady state nucleation rate J is constant beginning at the induction time τ leading to a volume fraction that is linear with time $v = (2a)^2 J(t-\tau)$, where a is the radius of a nucleating particle. For small volume fractions $v \ll 1$, this model gives

$$P \propto \exp\left(-\frac{L}{R}\sqrt{\frac{3}{2}[-1-\ln(4a^2 J(t-\tau))]}\right). \tag{2.15}$$

Our second model takes into account that structural disorder in a glass will make the nucleation barriers W random, leading to an exponential dispersion in nucleation times. Assuming the barrier distribution is uniform in some interval ΔW, the volume fraction will be logarithmic in time, [21]

$$v = \frac{kT}{\Delta W}\ln\frac{t}{\tau},$$

where τ corresponds to the minimum barrier in the system (assumed to be approxi-

mately the same as the induction time). For $v \ll 1$, this yields,

$$P \propto \exp\left[-\frac{L}{R}\sqrt{\frac{3}{2}\left(-1+\ln\frac{\Delta W}{kT\ln t/\tau}\right)}\right]. \tag{2.16}$$

The volume fraction with logarithmic time dependence can be well approximated by the Weibull distribution in a broad time interval prior to reaching the percolation threshold (Fig. 2-11), which is consistent with the experimental data.[7, 8, 10, 11] For numerical values, we use the data for the temperature dependent induction times and estimate the nucleation rates given in [19] for $Ge_2Sb_2Te_5$ as listed in Table 2.1. For the plots with a linearly time dependent volume fraction, a is set to 30 nm roughly estimating the larger diameter nuclei in that experiment and 0.1 is added to the volume fraction as an estimate for the volume fraction present at the induction time. The width of the nucleation barrier distribution ΔW remains an unknown parameter. Based on the estimates in Ref. [21] it is expected to be in the range of tenths of eV. Here, we rather arbitrarily set $\Delta W \sim 0.6$ eV reflecting the assumed condition $\Delta W \gg kT$.

The slope in the Weibull coordinates depends linearly on the thickness only. Increasing the area shifts the curves upward slightly without changing the slope while increasing the temperature shifts the curves to the left significantly (Fig. 2-12).

2.4.4 Activation Energy

In general, the activation energy is defined by $E_x \equiv |d(\ln x)/d(1/kT)|$ where x is the parameter being considered. Table 2.1 gives the numerical values for the activation energies of the crystal growth rate (E_u), nucleation rate (E_J), and induction time (E_τ) for $Ge_2Sb_2Te_5$ as obtained in Ref. [19].

An important data retention characteristic is the time required for the device to fail, referred to as the failure time. For each failure mechanism there is only one time-

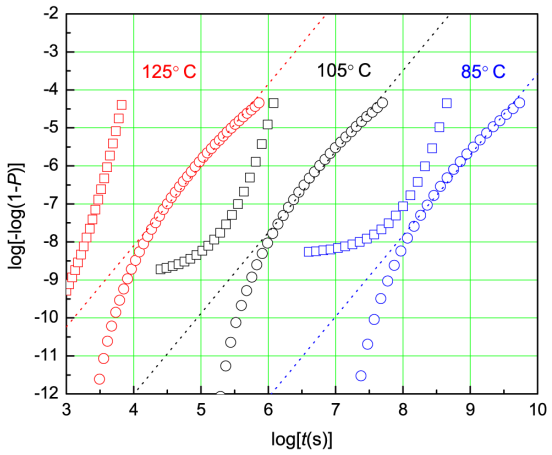

Figure 2-11: Weibull plots of failure statistics for the cases of fixed (□) and random (◯) nucleation barriers for various temperatures. Straight lines show linear approximation (Weibull law) in the range $v_0 < v < v_c$. The plots were generated using Eqs. (2.15) and (2.16) for a device with dimensions $L = 90$ nm and $A = 8100$ nm^2. Here log represents the base 10 logarithm.

dependent variable used to calculate the shunting probability. For the growth driven models of Sec. 2.4.1 and 2.4.2 it is the thickness $L(t) = L_0 - ut$ or the conducting particle radius $R(t) = R_0 + ut$, while for the nucleation models of Sec. 2.4.3 it is the volume fraction $v(t) = (2a)^2 J(t - \tau)$ or $v(t) = (kT/\Delta W) \ln t/\tau$ for the cases of fixed and randomly distributed nucleation barriers, respectively.

The activation energy for the failure time E_t for the different cases considered

Parameter (x)	$\ln(x_0)$	E_x (eV)
u (in pm/s)	72.0	2.35
J (in μm^{-2}s^{-1})	97.53	3.50
τ (in s)	-72.36	2.74

Table 2.1: Fitting parameters for the linear growth rate (u), nucleation rate (J), and induction time (τ) for $Ge_2Sb_2Te_5$ given in Ref. [19]. Data were fitted with an Arrhenius dependence $\ln(x) = \ln(x_0) \pm E_x/kT$.

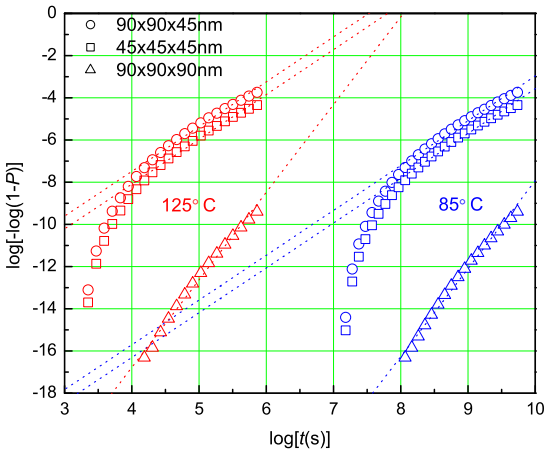

Figure 2-12: Weibull plots of failure statistics for devices with random nucleation barriers [Eq. (2.16)]and dimensions AxL. The dashed lines represent a linear approximation (Weibull law) where the volume fraction is restricted to the range $v_0 < v < v_c$; the curves at $v < v_0$ look qualitatively similar.

above can be obtained in a straight-forward manner. First, invert $L(t), R(t)$, or $v(t)$ to find the time t required to grow a fixed distance or nucleate a fixed number of particles. This time corresponds to the time required to obtain a specific (constant) value for the shunting probability. Then, using the definition of the activation energy, simply differentiate the logarithm of this time with respect to $1/kT$.

For example, to calculate the activation energy of the failure time for the case of randomly distributed barriers, we begin by inverting $v(t)$ to find $t = \tau \exp(v\Delta W/kT)$. The activation energy is then

$$E_t = \left| \frac{d(\ln \tau)}{d(1/kT)} + v\Delta W \right| = E_\tau + v\Delta W. \tag{2.17}$$

Notice that the above depends weakly on the volume fraction present when the activation energy is being measured since $E_\tau \gg v\Delta W$. Using the numerical values listed in Table 2.1 for E_τ and setting $\Delta W \sim 0.6$ eV, we find the activation energy to be

$E_t \sim 2.8 - 2.9$ eV for volume fractions under consideration $v \sim 0.1 - 0.3$.

Similarly, for both of the growth driven failure mechanisms $E_t = E_u \approx 2.35$ eV while for fixed-barrier nucleation $E_t = E_\tau + E_J \approx 5.85$ eV. Experimental data[10] suggest an activation energy of around $2.4 - 2.5$ eV. This implies that crystalline growth with little nucleation is most likely responsible for the failure of the cells in that study. Furthermore, the erroneously high activation energy predicted by employing fixed nucleation barriers suggest that this is not a suitable model for these materials. We would also like to point out that reliability studies involving different materials may have the opposite result depending on whether the material is dominated by nucleation or growth (determined by the relative activation energies).

2.5 Effect of a Weak External Electric Field

It is of interest to investigate the effect of an external field well below the threshold field E_{th} on the data retention characteristics. We consider the effect of an applied electric field on the evolution of an existing chain of conducting particles. Following the above, we assume that there are few other particles nearby and the chain is nearly linear and perpendicular to the bottom electrode as in Fig. 2-13. Due to depolarization, the partially completed shunt has the effect of increasing the electric field near its tip, similar to the behavior of a lightning rod. The field in the vicinity of the shunt is given by [20]

$$E = N^2 E_0 / \Lambda, \tag{2.18}$$

where E_0 is the value of applied field far from the chain and $\Lambda = \ln(2N) - 1$ is treated as approximately constant (for $2 < N < 40$, the relation $1 < \Lambda < 3$ holds).

The change in free energy due to the transformation of a spherical volume with radius r from the amorphous phase into the crystalline phase in the presence of the

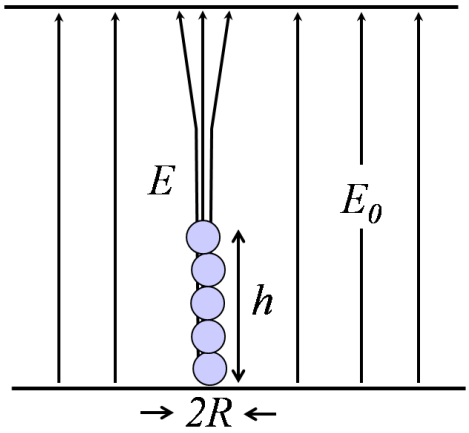

Figure 2-13: The effect of a weak external electric field E_0 on a partially formed shunt. The loose end of the chain concentrates the electric field, lowering the barrier for nucleation of particles in the nearby region.

applied field in Eq. 2.18 is given by

$$F = 4\pi r^2\sigma - \frac{4\pi}{3}\mu r^3 - \frac{\epsilon N^4 E_0^2}{2\Lambda^2} r^3, \tag{2.19}$$

where σ the surface tension, μ the chemical potential difference between the two phases, and ε is the dielectric permittivity of the host insulating phase. The first two terms represent the interface and bulk contributions (respectively) while the third term describes the change in electrostatic energy due to the presence of the nucleus. The free energy attains its maximum at

$$W = \frac{W_0}{(1 + N^4\zeta/4)^2}, \quad \zeta \equiv \frac{\varepsilon E_0^2 R^3}{\Lambda^2 W_0}, \tag{2.20}$$

where $W_0 \equiv 16\pi\sigma^3/3\mu^2$ is the classical result for the nucleation barrier in the absence of an applied field and $R = 2\sigma/\mu \sim 3$ nm is the zero field critical radius. Note that this is the same as the lattice spacing used throughout the paper. For fields below the threshold field E_{th} the parameter $\zeta \ll 1$. Using the typical values $W_0 \sim 3$ eV, $\varepsilon = 16$, and $E_{th} \sim 10^5$ V/cm, we have $\zeta < 0.1$. Below the threshold field (e.g. for

read fields) $\zeta \ll 0.1$

The nucleation barrier corresponds to the radius $r = R/(1+N^4\zeta/4)$. For relatively small chains or weak enough fields, $r \sim R$. That is, the size of a particle nucleating under weak field conditions will be approximately the same as the classical case of zero field. By expanding the denominator of Eq. (2.20) in a Taylor series we have

$$W = W_0 \left(1 - N^4\zeta/2\right) \tag{2.21}$$

The nucleation rate will be substantially increased when $W - W_0 \gtrsim kT$. From Eq. (2.21), this implies that

$$N \gtrsim \left(\frac{2kT}{\zeta W_0}\right) = \left(\frac{2\Lambda^2 kT}{\varepsilon E_0^2 R^3}\right)^{1/4}. \tag{2.22}$$

Eq. (2.22) gives the number of particles in a nearly linear chain that are required in order to lower the nucleation barrier significantly enough so that a shunt will quickly form. Once one particle has nucleated near the tip, subsequent nucleation is further enhanced. For the numerical values above along with $\Lambda \sim 2$, $kT = 0.03$eV, and $E_0 \sim 10^4$ V/cm we have $N \gtrsim 7$. Notice that the required number of particles is $N \propto E_0^{-1/2}$. By setting $N = L/2R$, we find the field required for such premature shunting to occur is

$$E_0 \gtrsim \left(\frac{32R\Lambda^2 kT}{\varepsilon L^4}\right)^{1/2} \tag{2.23}$$

We observe that the field required for a significant change in nucleation to occur increases with temperature. This is due to the condition that the lowering of the barrier should be much greater than the thermal limit kT. For stronger fields, the length of a conducting chain required to shunt the device can be substantially less than the device thickness L.

Using the same numerical values as before and setting $L/2R = 10$ (corresponding to a 60 nm thick device) we have $E_0 \gtrsim 10^3$ V/cm. We note that the latter prediction would not be applicable to the case of the read-out field that exists over rather short times ($\sim$ ns) insufficient for nucleation to occur consistently. For stronger fields that are still below the threshold, the rate of nucleation increases significantly enough so that even such short voltage pulses are capable of affecting the retention ability.

2.6 Conclusions

In conclusion, we have developed a theory of crystalline path formation in thin, insulating layers wherein the crystalline volume fraction is lower than the percolation threshold. In particular, we have examined the implications of our model with respect to the data retention characteristics of chalcogenide phase change memory devices. The results of the theory are summarized below.

1) In thin film structures, shunting is due to conductive paths that are nearly rectilinear and the number of particles that form those paths is expressed analytically in terms of device and material parameters. Our model predicts that the formation probability of such paths increases with increasing volume fraction and increases exponentially with decreasing device thickness. Shunting probability is also directly proportional to the device area. The critical thickness and area below which rectilinear paths dominate over percolation clusters have been derived. Critical device dimensions were extended to include large arrays, showing that our model applies to modern devices comprised of $\sim 10^9$ cells and thicknesses of less than 100 nm.

2) Our numerical simulations independently support our analytical results regarding the thickness and area dependencies of the shunting probability, as well as the rectilinear nature of the crystalline paths when the volume fraction is below the percolation threshold.

3) The effects of various phase change memory device geometries were examined and it was determined that the relevant length to consider for any geometry is the minimum distance between the bottom electrode and surrounding crystalline layer. Our analytical expression for the shunting probability is simply modified to reflect that fact for any geometry considered.

4) There are a number of different mechanisms through which the failure probability varies with temperature and evolves over time. We have considered several mechanisms where the conductive phase volume fraction increases over time, including: (a) decaying volume of the insulating amorphous region due to peripheral crystallization; (b) growth of individual crystalline particles; and (c) nucleation of new crystalline particle in a glass system of: (i) constant nucleation barriers, and (ii) uniformly distributed random nucleation barriers. For each mechanism, the time and temperature dependence of the failure probability was derived and the corresponding characteristic Weibull plots of the failure statistics were generated. Activation energies for the failure time related to each mechanism were derived and compared to available experimental data. Our results suggest that data retention failure can be due to either crystal growth or nucleation (with dispersed barriers) depending on the material being considered.

5) An external electric field effects the probability of shunt formation by lowering the nucleation barriers in the vicinity of a partially completed shunt and by favoring linear chains of conducting particles.

Chapter 3

Thermodynamics of Conductive Filaments in Threshold Switches

Chalcogenide glasses exhibit reversible switching between highly resistive (amorphous) and conductive (crystalline) phases when subjected to appropriate voltage pulses. This phenomenon recently regained interest in connection with phase change memory applications. [22] Another application is found with threshold switches (TS), which require a minimum holding voltage or current to sustain the conductive state. [23] A new type of chalcogenide devices combines TS with phase change memory.[45]

It is known that upon switching, a high-current filament forms,[24] the radius of which increases with current as [25] $r \propto I^{1/2}$. This relation is often cited but remains poorly understood; at present, there is no theory relating the filament properties to material parameters. An approach based on the principle of least entropy production [26] did not lead to specific predictions. The validity of that principle remains questionable, [27] and avoidance of it leads to different results, as shown here.

In this chapter we introduce a thermodynamic theory of steady state conductive filaments. It predicts the filament radius vs. the electric current and material parameters as well as the corresponding current-voltage (IV) characteristics. A finite element numerical model is employed to support our analytical results.

3.1 Theoretical Framework

Whether electronic[29, 28] or crystalline,[30] the conductive filament represents a domain of different phase in the insulating host, thus calling upon the analysis of phase equilibrium. Our conservative approach avoids the principle of least entropy production starting instead with the kinetic Fokker-Planck equation (see e.g. Ref. [31], p. 428) in the space of cylinder radii r,

$$\frac{\partial f}{\partial t} = -\frac{\partial s}{\partial r}, \quad s \equiv -B\frac{\partial f}{\partial r} + Af = -Bf_0\frac{\partial}{\partial r}\left(\frac{f}{f_0}\right). \tag{3.1}$$

Here, f is the distribution function so that $f(r)dr$ gives the concentration of filaments in the interval $(r, r+dr)$; s is the flux in radii space (s^{-1} cm^{-3}). B is the 'filament radius diffusion coefficient'; A is connected with B by a relationship which follows from the fact that $s = 0$ for the equilibrium distribution $f_0(r) \propto \exp[-F(r)/kT]$, where F is the free energy, k is Boltzmann's constant, and T is the temperature. Applying the boundary condition $f(r = 0) = 0$ to Eq. (3.1) reflects the fact that very thin filaments cannot exist due to limitations such as loss of conductivity or mechanical instability (extraneous to the present model). Another condition $f(r = \infty) = 0$ implies finite radii achievable over finite times t.

Using the right-hand-side expression for s, multiplying Eq. (3.1) by r, integrating from 0 to ∞ by parts , and noting that $\int frdr = \langle r\rangle$, yields $\partial\langle r\rangle/\partial t = \langle \partial F/\partial r\rangle$. We then approximate $\langle F\rangle = F(\langle r\rangle)$ and $\langle \partial F/\partial r\rangle = \partial\langle F\rangle/\partial\langle r\rangle$, thereby neglecting fluctuations in the ensemble of nominally identical filaments. Omitting for brevity the angular brackets, one finally obtains,

$$\frac{\partial r}{\partial t} = -b\frac{\partial F}{\partial r} \quad \text{with} \quad b = \frac{B}{kT}. \tag{3.2}$$

This equation has the standard meaning of a relation between the (growth) velocity and the (thermodynamic) force $-\partial F/\partial r$ with the mobility b and the diffusion

coefficient B obeying the Einstein relation.

While Eq. (3.2) can, in principle, describe transient behavior, here we limit ourselves to the steady state case $\partial r/\partial t = 0$, which, according to Eq. (3.2) takes place when the free energy is a minimum (obviously different from the condition of least entropy production [26]). While adequate for understanding the filament, this model is lacking the feature of blocking electrodes typical of TS, [25, 29] which can cause deviations between the predicted and measured quantities.

3.2 The Free Energy of a Conductive Filament

The general form for the free energy of a system containing a cylindrical filament of radius r and height h is given by

$$F = \int C_v \delta T dx^3 + \frac{1}{8\pi} \int \varepsilon \left|\mathbf{E}\right|^2 dx^3 + 2\pi r h \sigma + \pi r^2 h \mu. \tag{3.3}$$

where the integration is taken over the entire device volume. Here, C_v is the volumetric heat capacity, δT is the current dependent temperature change, $\mathbf{E}$ is the electrostatic field within the device, ε is the dielectric permittivity, σ is the surface energy, and μ is the change in chemical potential between the two phases. Correspondingly, the first term in Eq. (3.3) represents the thermal contribution, the second one stands for the electric energy, and the last two terms correspond to the phase transformation.

A positive sign for μ reflects the fact that the system is stable under zero bias, i.e. a conductive filament will not spontaneously form under such conditions. Furthermore, upon field removal a fully formed filament will decrease in size as dictated by Eq. (3.2) and eventually disappear. It is this characteristic of threshold switches which leads to the requirement of a holding current. It has been observed experimentally that there is a maximum time in which the holding bias may be removed before the device

switches back into the OFF state, referred to as the maximum interruption time $\tau_{\max}$. Once this occurs, it is necessary to reapply the threshold voltage to return the device to the ON state. Future work will apply the framework of this model in order to calculate $\tau_{\max}$ as a function of current and material parameters.

3.2.1 Analytical Model

To analytically describe the free energy we consider a model based on a flat plate capacitor of area A and thickness h containing a conductive cylindrical filament of radius r, as shown in Fig. 3-1.

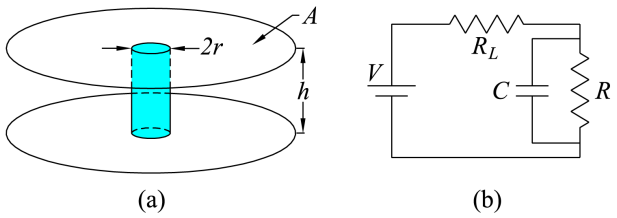

Figure 3-1: (a) the model capacitor structure of area A, height h, and filament radius r. (b) circuit schematic showing source voltage V, load resistance R_L, capacitance C, and filament resistance R.

Experimental data[25] suggest that the filament radius is greater than 1 μm for all operating currents. Assuming the device is sufficiently thin ($h \lesssim r$), the current density will be nearly uniform and concentrated within the filament. As such, the temperature increase due to Joule heating occurs mostly within the filament and the integral for the thermal contribution may be approximated by the average temperature increase within the filament ΔT multiplied by the filament volume. Similarly, since the field within a parallel plate capacitor is approximately uniform and perpendicular to the surface, the integral for the electrostatic contribution is proportional to the magnitude of the field squared. From Eq. (3.3) the free energy is then given

by,

$$F = C_v \Delta T \pi r^2 h + \frac{E^2 \varepsilon}{8\pi} A h + 2\pi r h \sigma + \pi r^2 h \mu. \tag{3.4}$$

The field and temperature in Eq. (3.4) are

$$E = \frac{V}{h} \frac{R}{R + R_L}, \quad R = \frac{\rho h}{\pi r^2}, \quad \Delta T = \frac{I^2 h^2 \rho}{8\pi^2 \chi r^4}, \tag{3.5}$$

where V is the source voltage, R_L is the load resistance, I is the current, and R and ρ are the filament resistance and resistivity, respectively. χ is the thermal conductivity taken to be the same for the filament and host materials. In Appendix B the above equations are derived along with their corresponding range of applicability.

Substituting Eqs. (3.5) into Eq. (3.4) the free energy becomes

$$F = \frac{3Wh}{2r_0} \left\{ \frac{\beta x^2}{(1 + Hx^2)^2} + \frac{\gamma}{(1 + Hx^2)^2} + x + x^2 \right\}, \tag{3.6}$$

where $x \equiv r/r_0$ and we have introduced the dimensionless parameters

$$\beta = \frac{\pi r_0^3 V^2}{12 W \kappa \rho}, \quad \gamma = \frac{r_0}{h} \frac{C V^2}{3W}, \quad H = \frac{R_L \pi r_0^2}{\rho h}, \tag{3.7}$$

with $\kappa = \chi / C_v$ being the thermal diffusivity. Here, W and r_0 are the parameters of energy and length defined as

$$W = 16\pi\sigma^3/3\mu^2 \quad \text{and} \quad r_0 = 2\sigma/\mu. \tag{3.8}$$

They would have the physical meaning of nucleation barrier and radius in classical nucleation theory (in which μ is negative and $|\mu|$ is used instead). Assuming σ and $|\mu|$ to be of the same order of magnitude as for crystal nucleation in chalcogenide glasses, one can use the corresponding estimates[32] $W \sim 2$ eV and $r_0 \sim 3$ nm.

Numerical values of other parameters in Eq. (3.7) are as follows: $V \sim 0.1 - 1$ V,

$\rho \sim 0.001 - 0.1\ \Omega\cdot$cm, $\kappa \sim 10^{-3}$ cm^{2}s^{-1}, $h \sim 30 - 3000$ nm, $R_L \sim 10 - 100\ \Omega$, $\varepsilon \sim 10$, and $A \sim 10^4 - 10^{10}$ nm^2 (with $C = A\varepsilon/4\pi h$). The lower dimensions correspond to modern devices [45], while the larger ones are typical of earlier structures. [24, 25, 23, 33]

In what follows we concentrate on the latter because of the availability of published experimental results. Given the typical values of the corresponding parameters, one can neglect the second and the third terms in Eq. (3.6), which allows analytical solutions.

3.2.2 Operating Characteristics in Terms of Device Parameters

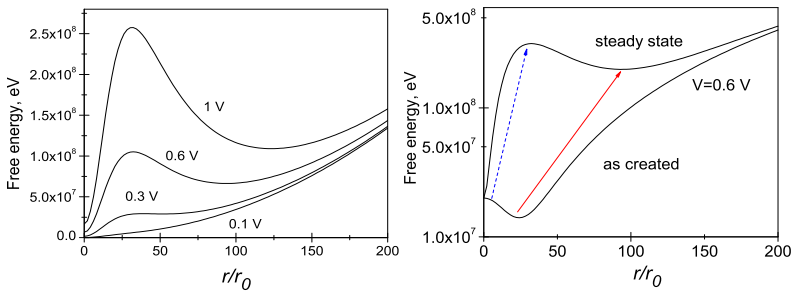

Figure 3-2: Left: Free energy of a structure with conductive filament as described by Eq. (3.6). Right: Free energy of a filament as created (thermal contribution neglected) and in the steady state. Arrows show how the energy minimum moves to the right and becomes metastable separated by a barrier from the state without the filament.

As illustrated in Fig. 3-2, the shapes of the free energy curves corresponding to the parameters of older era devices indicate that the filament can exist in a metastable state [i.e. with the right minimum (at finite x) shallower than the left one (at $x = 0$)]. It becomes stable under unrealistically high voltages $V > (h^2/CR_L)\sqrt{3W\rho/\pi r_0^3\kappa} \sim$

300 V. On the other hand, finite radius filaments become unstable under applied voltages below,

$$V_0 = 18\sqrt{W\kappa\rho/\pi r_0^3} \sim 0.1 \text{ V}. \tag{3.9}$$

V_0 is defined by the conditions $\partial F/\partial r = \partial^2 F/\partial r^2 = 0$ and is presented by the curve labeled 0.3 V in Fig. 3-2. The minimum filament radius at $V = V_0$ is

$$r_{min} = r_0\sqrt{2/H} \sim 10^3 r_0 \sim 1 \ \mu\text{m}.$$

Its related filament resistance is $R_{max} = \rho h/\pi r_{min}^2 = R_L/2$. Using the common terminology for the minimum voltage across the device under which the filament can exist, $V_h = V_0/3$ is the holding voltage (keeping in the voltage across the device is given in terms of the applied voltage by $V_d = V_a/(1+R_L/R)$ as discussed in Appendix B. One can define the holding current as,

$$I_h = V_0/(R_{max} + R_L) = 2V_0/3R_L \sim 10 \text{ mA}. \tag{3.10}$$

While metastable, the filament is predicted to be extremely long-lived under source voltages just slightly above V_0. Indeed, as seen from Fig. 3-2, the barrier separating the metastable minimum can be as high as $W_B \sim 10^8$ eV; hence, an extremely large exponential of its thermal activation, $\exp(W_B/kT)$, making the issue of metastability practically immaterial. More quantitatively, the shape of the free energy at voltages close to V_0 is described by the expansion $\delta F = (1/2!)(\partial^2 F/\partial x \partial V)|_{V_0,x_{min}}\delta x \delta V + (1/3!)(\partial^3 F/\partial x^3)|_{V_0,x_{min}}(\delta x)^3$ where $x_{min} = r_{min}/r_0$, $\delta x = x - x_{min}$ and $\delta V = V - V_0$ yielding

$$r = r_{min}\left(1 + \sqrt{\frac{\delta V}{4V_0}}\right), W_B = \frac{4hW}{r_0 H}\left(\frac{\delta V}{V_0}\right)^{3/2}. \tag{3.11}$$

The large barrier values are due to a large number of particles constituting the filament: $h/Hr_0 \sim hr_{min}^2/r_0^3 \gg 1$.

The metastable nature of a steady state filament does not appear with the filament immediately upon creation, when the thermal contribution is not present [i.e. the first term is excluded from Eqs. (3.4) and (3.6)]. Such a newly created filament is stable under voltages exceeding

$$V_c = \frac{h}{r_0}\sqrt{\frac{6W\rho}{CR_L r_0}} \sim 30 \text{ V}. \tag{3.12}$$

While its meaning is similar to that of threshold voltage, the two have different values due to the limitations of the present model that considers filament creation and disappearance in one step processes, neglecting the possibility of filament nucleation. [30]

Beyond the critical region of $I - I_h \ll I_h$, by differentiating Eq. (3.6) with respect to r and setting equal to zero the stable filament radius is found to be

$$r = r_0 \left(\frac{\rho h^2}{12\pi\kappa W r_0}\right)^{1/4} \sqrt{I} \quad \text{when} \quad I - I_h \gg I_h, \tag{3.13}$$

which is proportional to $\sqrt{I}$, consistent with the above mentioned experimental observations. It predicts the current voltage characteristics, in which the device voltage $V_d = IR \propto I/r^2$ remains constant. Its value extrapolated from the the graph of the high current $I \gg I_h$ parallel to the vertical axis is given by

$$V_{h\infty} = V_h/\sqrt{3} = \sqrt{12W\kappa\rho/\pi r_0^3} \tag{3.14}$$

Experimentally, the infinitely steep current graph can be masked by the effects of electrodes. The inequality $V_{h\infty} < V_h$ implies a 'knee' in IV characteristics which is a region of negative differential resistance. The current density within the filament (given by $J = I/\pi r^2$) depends on material parameters and device geometry, and is

independent of current,

$$J = \sqrt{\frac{12\kappa W}{\pi \rho r_0^3 h^2}}, \tag{3.15}$$

until filament diameter reaches the device size, after which it linearly increases with current. Using the parameters from Ref. [25], Eq. (3.15) predicts $J \sim 10^4$ A/cm^2 in agreement with the data. For the modern devices, the characteristic current density can reach $J \sim 10^7$ A/cm^2.

3.2.3 Numerical Simulation

Using the COMSOL multiphysics package we have performed numerical simulations of a specific real structure (Fig. 3-3) described in Ref. [25]. These devices are fabricated by applying a layer of insulating dielectric material (such as SiO_2) over an electrical contract (such as Molybdenum), then a layer of chalcogenide is rf-sputtered with a target composition of $Te_{39}As_{36}Si_{17}Ge_7P_1$. Lastly, the structure is capped with an aluminum contact pad. By applying a sharp dent to the top electrode (not shown in the figure), a region is formed within the chalcogenide material which is relatively thinner than its surroundings. As such, this region will have a lower resistance. By then applying a very large current, a pore is formed within the dielectric layer. Since the dielectric is relatively insulating, the current is focused within the active region dictated by the pore diameter.

COMSOL uses the finite element method (FEM) to solve Laplace's's equation and the heat conduction equation simultaneously, coupled via Joule heating

$$\nabla \cdot \mathbf{J} = 0, \quad \nabla \cdot \chi \nabla T = -\rho \left|\mathbf{J}\right|^2 \tag{3.16}$$

This allows for the free energy to be calculated for arbitrary geometries and also provides a check for the approximations used in the analytical modeling as discussed in Sec. 3.2.4. The electric field and temperature distributions are calculated within

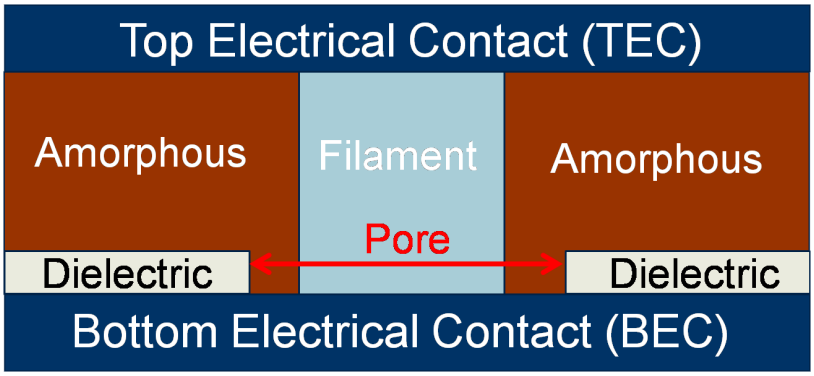

Figure 3-3: TS device described in Ref. [25] used in the numerical modeling. Device dimensions are as follows: device radius: 100-125 μm, pore radius: 5-50 μm, chalcogenide thickness: 1-2 μm, electrode thickness: 1 μm, and dielectric thickness: 0.3 μm

the device as dictated by the geometry given in Fig. 3-3 and material parameters given in Table 3.1. These quantities are then used to evaluate the integrals in Eq. (3.3) for a given filament radius. COMSOL also allows for the material properties within each region to depend on temperature and voltage. This feature was used to provide the electrical conductivity of the crystalline filament an Arrhenius temperature dependence with activation energy given in Ref. [34].

Material (x)	χ (W/m·K	ρ (Ω·cm)
filament	0.5	see text
amorphous	0.17	33.0
electrode	25	2.6×10^{-6}
dielectric	1.4	10^{18}

Table 3.1: Material parameters used in the numerical modeling. All properties are assumed isotropic within each region except for the electrical conductivity of the conductive filament, which is assumed to have an Arrhenius temperature dependence.

To find the filament radius which minimizes the free energy, the MATLAB interface to COMSOL was employed. This allows the rich data-processing environment of MATLAB to access the COMSOL solver through a series of scripts (referred to as *m-files*). The program consists of two major scripts: 'COMSOL.m' and 'FindMin.m'.

The first is used to call the COMSOL solver to calculate the electric field and temperature distribution for some fixed filament radius. The second performs a search for the radius which minimizes the free energy. It starts by calculating the free energy for some arbitrarily small filament radius (typically the nucleation radius ~ 3 nm is used) and then iterates the radius until it reaches some maximum size (usually the size of the device). This provides a map of the free energy as a function of filament radius. Lastly, it searches for the radius which has the lowest corresponding free energy. The program flow is illustrated in Fig. (3-4).

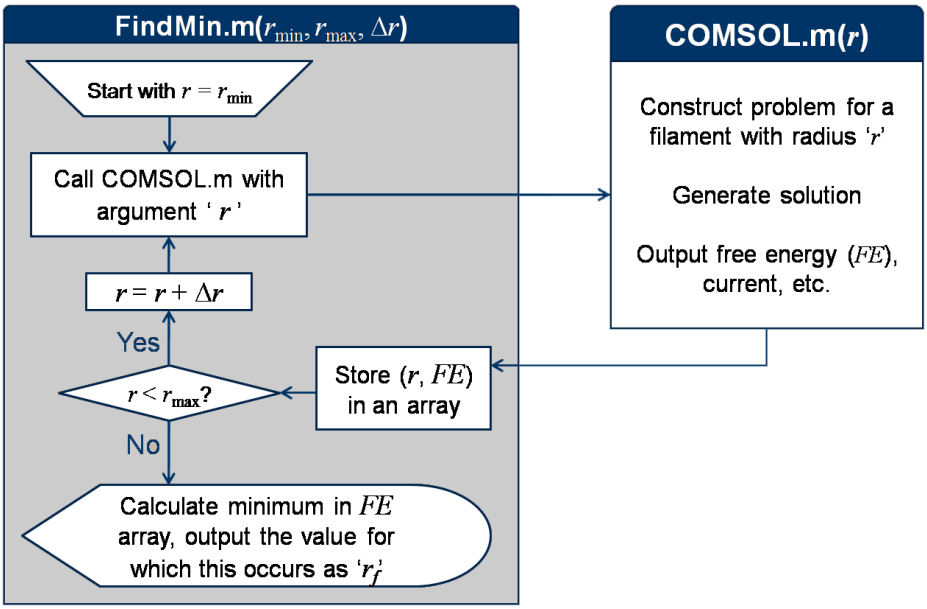

Figure 3-4: Program flow for the MATLAB scripts used to find the filament radius which minimizes the free energy.

In order to keep the number of simulations low, at first a relatively coarse step size is used. Each successive run uses the previous minimum and begins the search one step to the left and ends one step to the right using progressively finer stepping as illustrated in Fig. (3-5). The process is repeated until the desired accuracy is achieved. As an example, for a device with a pore radius of 10 μm, in order to calculate the critical filament radius to within 10 nm, 1000 iterations would need to

be performed, each taking only a few seconds. Using a search algorithm which "hones in" on the critical radius reduces this to only 50 iterations. The benefit is evident when you consider that finding one critical radius only corresponds to one data point on an IV curve. Generating an IV curve with 10 data points takes only $\sim$ 10 minutes with this method as opposed to 3 hours.

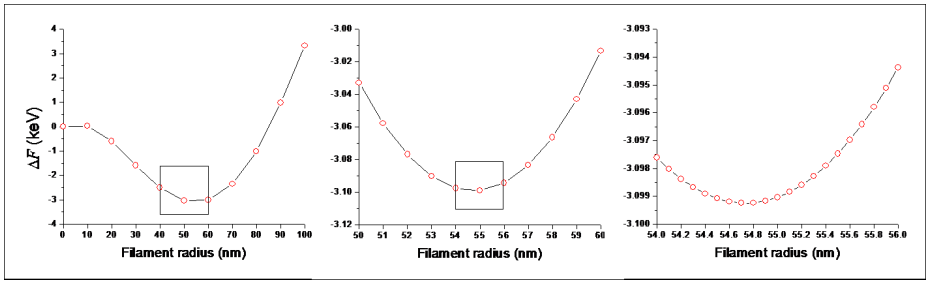

Figure 3-5: Illustration of algorithm used to find the filament radius by using a progressively finer stepping.

3.2.4 Verification of Approximations Used in the Analytical Model

In addition to being able to simulate complicated device structures that are not amenable to an analytical treatment, the use of COMSOL also allows us to verify that the approximations used in the analytical modeling are appropriate. This provides a check for the error involved and an estimate for their range of applicability.

In Eq. (3.4) the thermal contribution to the free energy is approximated as being the average temperature increase within the filament multiplied by its volume. The more general form in Eq. (3.3) includes the contributions from surrounding material. COMSOL was used to simulate the temperature distribution for a device with a radius of 10 μm and thickness 2 μm and the thermal contribution to the free energy was calculated within each region. As illustrated in Fig. 3-6, the thermal term due to the filament alone is well approximated by using it's average temperature [Eq. (B.5)]

while the total contribution (which includes the surrounding amorphous material) is better approximated by the maximum temperature increase [Eq. (3.5)] within the filament.

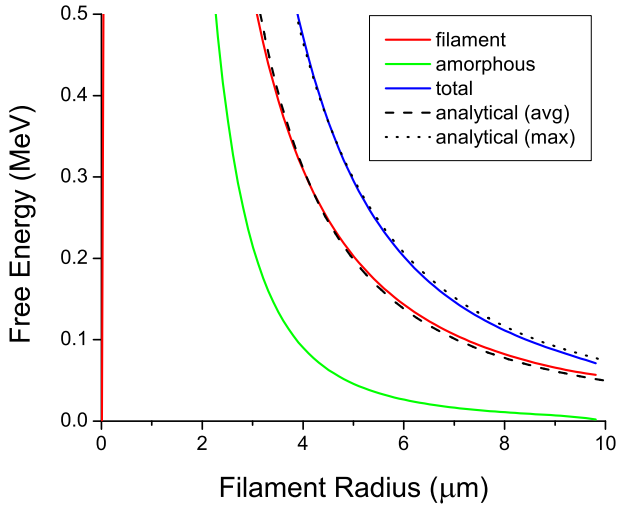

Figure 3-6: Comparison of the thermal contribution to the free energy as a function of filament radius. The average filament temperature is given by Eq. (B.5) and the maximum filament temperature is given by Eq. (3.5).

A similar comparison was made for the voltage across the device. The approximation used in Eq. (3.5) is in good agreement except for the case of extremely narrow filaments. As discussed in Appendix B, the resistance given in Eq. (3.5) is valid when $r^2 \gg A\rho_c/\rho_a \approx 10^{-4}$ A using standard values for the resistivities (see Table 3.1). For a device whose area A is on the order of 100 μm^2, this implies the filament radius needs to be $\gtrsim 100$ nm for the approximation to hold as can be seen in Fig. 3-7

3.3 Comparison with Experimental Data

Remarkably good agreement was obtained for the filament radius vs. current, as illustrated in Fig. 3-8. It is important to note that it was not necessary to adjust

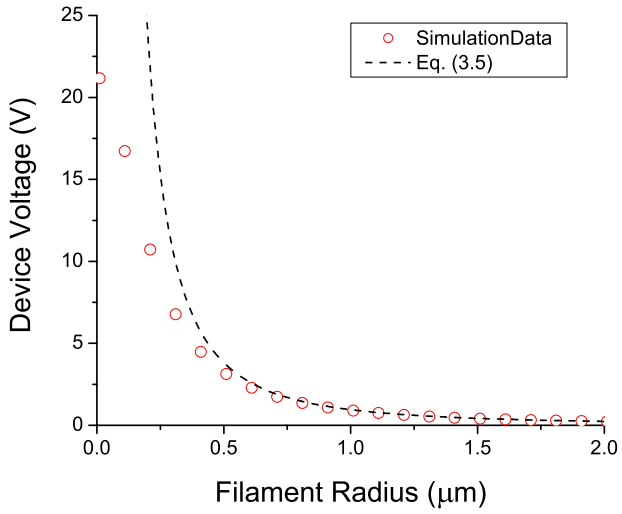

Figure 3-7: The simulated voltage across a TS device as a function of filament radius compared to the approximation used in Eq. (3.5). Note how the approximation becomes less accurate for extremley narrow $\sim .1$ μm filaments.

the material parameters used in the simulations in order to fit the data, instead the standard values given in Table 3.1 were used. The experimental data extracted from Ref. [25] is fairly scattered showing little dependence on thickness or pore diameter. In those experiments, devices with thicknesses ranging from 1.5-2 μm and pore diameters between 6-25 μm were used (see Fig. 3-3) and the simulations used a device with thickness 2μm and pore diameters of 10 and 25 μm.

The current density was found to be nearly independent of current until the filament grew to the pore diameter limiting the current flow. Its value before saturation, the current for which saturation occurs, and the slope after saturation are all in reasonably good agreement with the experimental. In Fig. 3-9 a comparison between the analytical result in Eq. (3.15) and the data in Ref. [25] is made. Simulations were performed for two different pore diameters in order to illustrate the effects of saturation.

For sufficiently large currents ($I \gg I_h$), the voltage found across the device is con-

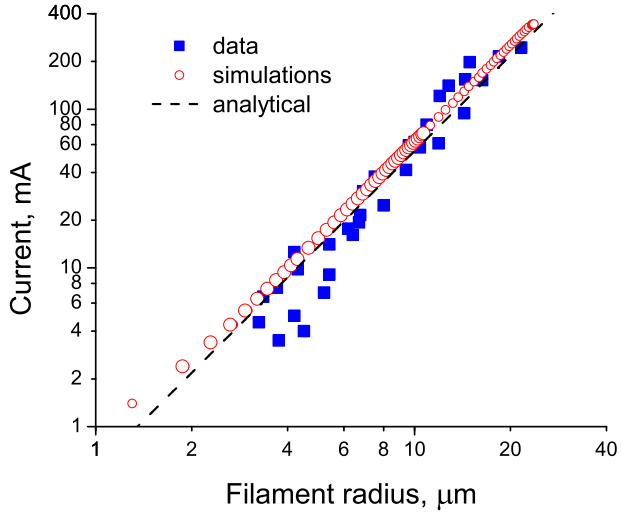

Figure 3-8: Filament radius as a function of applied current. Blue boxes are experimental data extracted from Ref. [25], red circles is data generated using the numerical simulation described in Sec. 3.2.3, and the dashed line is the analytical result given by Eq. (3.13).

stant and is given by Eq. (3.14). As you decrease the current, the voltage across the device increases until it reaches a maximum value at the holding current I_h. Decreasing the current further will cause the device to switch back into the OFF (amorphous) state. As a consequence, the region $I > I_h$ is that of negative differential resistance (NDR) and has been observed as a 'knee' in the current voltage characteristics. Our modeling predicts such behavior as indicated in Fig. 3-10

The simulated IV characteristic in Fig. 3-10 was limited to the filament (rather than to the entire device) since the blocking electrodes have not been taken into account. Our current level of understanding [23, 35] suggests that adding those electrodes to the model will shift the entire IV curve to the right by approximately 1 V, thereby making it very similar to the observed IV curves. [25]

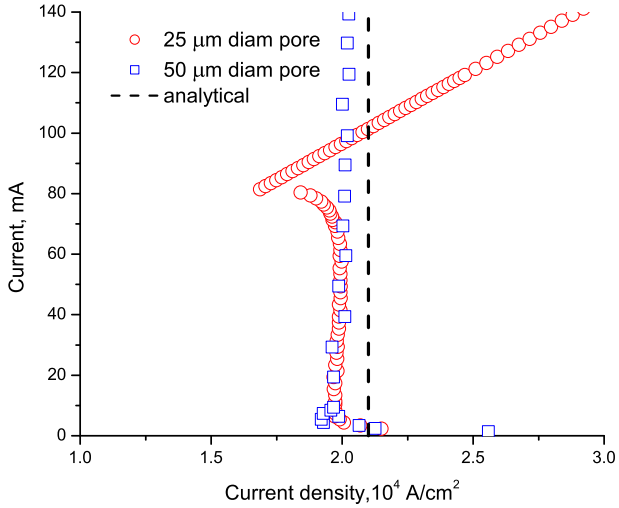

Figure 3-9: Current density in a conductive filament versus applied current. the devices simulated had two different pore diameters and the dashed line is the analytical result given by Eq. (3.15).

3.4 Conclusions

In summary, we have developed a thermodynamic theory of steady state conductive filament starting from the basic kinetic approach (Fokker-Planck equation). It predicts the filament properties observed in threshold switches: the existence of holding voltage, filament radius vs. current, and the shape of IV characteristics. A numerical routine was developed using the COMSOL and MATLAB software packages which can calculate the free energy for a device with arbitrary geometry and find the corresponding stable filament radius and current voltage characteristics. Future work will extend this theory to include the effects of blocking electrodes and a transient analysis.

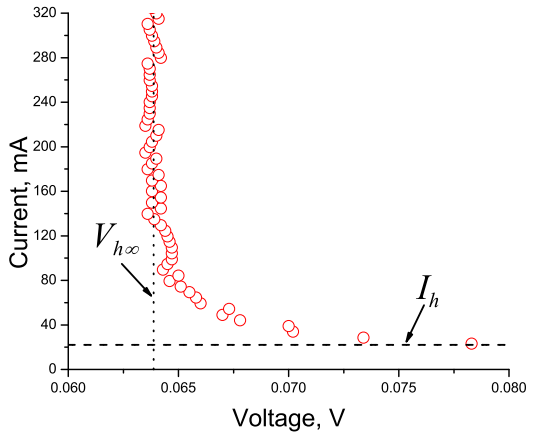

Figure 3-10: Simulated IV characteristics of a TS device. The holding current (I_h) is given by Eq. (3.10) and the minimum voltage across the device ($V_{h\infty}$) by Eq. (3.14).

Chapter 4

Charge Transport in Chalcogenide Glasses of Phase Change Memory

Chalcogenide materials have recently regained strong interest due to their ability to repeatedly transform between glassy (disordered) and crystalline (ordered) atomic structures. One application is the storage of digital data where 1s and 0s are recorded as either glassy (high resistive and low-reflective) or crystalline (low resistive and high reflective) structures. For example, optical memory disks use laser light to convert small portions of a thin chalcogenide film between the high and low reflective states. On the other hand, phase change memory (PCM) uses a voltage bias to convert the material between the high and low resistive states. PCM stores data in a smaller area and with higher speeds for both read and write processes than the optical memory disks.

PCM is an emerging nonvolatile memory technology with the capability of random access memory, it is sometimes referred as unified memory. Applications explored for this technology span from wireless, embedded systems [36] to solid state storage, [37] automotive, [38] and space applications. [39] Most recently, usage of PCM in computer applications was suggested as Storage Class Memory (SCM). [40]

Large, up to 1 gigabyte, memory arrays with PCM elements have been demonstrated for 180 nm, [41] 90 nm, [42, 43] and 45 nm [44] technology nodes. In PCM,

each individual element is in series with an access/selector device. Both MOS-based [41] and BJT/diode-based [43] selectors have been integrated with PCM.

Recently, PCM was integrated with a chalcogenide based thin film selector to form PCMS arrays, [45] opening a path for 3D stackable cross point phase change memory. Understanding and optimizing the material properties of chalcogenide nanoglasses in PCMS cells is a key enabler for this promising nonvolatile memory technology. [46]

The operation of PCM depends on charge transport in their constituent inclusions of chalcogenide glasses. When the device is in the reset state, the electrical conduction can be non-ohmic under practical voltages and temperatures. This non-ohmicity provides a way of supplying energy to the device faster than ohmic conduction and it needs to be properly understood in order to improve future device parameters.

The goal of this chapter is to recall the established physics of chalcogenide glasses and convey a broad picture of different mechanisms that are relevant to the problem of non-ohmic conduction in these materials. Ultimately, we provide a starting point for the additional studies that are required to better understand charge transport in PCM glasses.

The commonly observed nonlinear current-voltage (IV) characteristics (above $\sim 10^3 - 10^4$ V/cm) are often attributed to the Poole-Frenkel (PF) effect after the classical work [47, 48, 49] suggesting their plausible interpretation. An experimental signature of PF conduction is a region of linearity in the plot of $\ln(I/I_0)$ vs. either $\sqrt{V}$ or V where I_0 is the pre-exponential factor. The underlying mechanism is commonly related to the field-induced increase in free carrier concentration, as reflected in Refs. [22, 50, 51, 52, 53, 54, 55, 18] (except Ref. [56], which proposes hopping conduction).

Although there is general agreement about the observed PF-type of non-ohmicity and the fact that $I_0 \propto \exp(-E_a/kT)$, where E_a is the activation energy, k is the Boltzmann constant, and T is temperature, particular features observed and especially their interpretations vary dramatically between researchers. We note, for example,

that Refs. [22, 50, 51, 52, 18] present their observed non-ohmicity as $\ln(I/I_0) \propto \sqrt{V}$, while Refs. [57, 58, 53, 55, 56, 59, 60] describe their observations as $\ln(I/I_0) \propto V$. Furthermore, some of the latter results [57, 60] point at two different domains in the IV data which exhibit different proportionality coefficients and temperature dependencies.

In this chapter we frame what is known about d.c. conduction in chalcogenide glasses, indicate shortcomings in our current state of understanding, and suggest avenues for further investigation. We begin with a brief overview of the pertinent experimental data to provide some context for the key observations. That is followed by a review of the physics of localized states that underlies the unique properties of chalcogenide glasses. Then we provide a survey of conduction mechanisms that may explain the observed non-ohmic IV data, including: 1) the original Poole-Frenkel mechanism; 2) Schottky decrease in interfacial barrier near device electrodes; 3) field-induced delocalization of shallow band tail states near the mobility edges; 4) space charge limited (injection) currents; and 5) field effects in hopping conduction. Here, we discuss these and some other possible mechanisms of d.c. conduction in chalcogenide glasses including bulk materials and thin films down to the nanometer scale. Finally, we provide a summary of the candidate mechanisms and discuss their validity and implications, along with new indicative facts that are required to further evaluate these mechanisms.

4.1 Experimental Data: A Brief Overview

The interpretation of experimental data related to d.c. conduction in amorphous chalcogenides must consider fabrication technology, whether the sample is amorphous or vitreous in nature, cell geometry, and other factors. These materials have been intensely investigated over the past half century, but in this brief overview we present only some representative results for bulk and thin glassy chalcogenides with various

compositions, including the typical composition of $Ge_2Sb_2Te_5$ (GST) for modern PCM devices.

Certain universal features can be cited for amorphous chalcogenides, including the thermally activated conductivity $\sigma \propto \exp(-E_a/kT)$, positive thermopower indicative of p-type conduction, and negative Hall coefficient.[35] In magnitude, the activation energy for conduction, E_a, is close to half the mobility gap and can range from 0.3 to 1 eV, with a typical value of $E_a \sim 0.37$ eV for modern PCM devices under low bias.[61] Hopping conductivity with its classic temperature dependence, $\sigma \propto \exp[(-T_0/T)^{1/4}]$, is generally not observed, with the exception[62] of some unannealed, sputtered films.

Historically,[33] the study of bulk chalcogenide glasses revealed d.c. conductivity described by $\sigma \propto \exp(F/F_0)$ for fields up to $F \sim 1 - 4 \times 10^5$ V/cm, above which there was a steep increase in the field dependence. Typically, there was also an ohmic region observed at fields below $10^3 - 10^4$ V/cm; that transition field was found to increase linearly with thickness.[60] In some cases, usually below room temperature, two distinct exponential regimes were observed: a lower field region with $\ln\sigma \propto (F/F_{01})$ followed by a steeper region with $\ln\sigma \propto (F/F_{02})$. [60, 63, 64] The slopes F_{01} and F_{02} had opposite temperature dependencies and F_{01} was independent of sample thickness while F_{02} increased linearly with thickness (indicative of space charge limited current, as discussed in Sec. 4.3.4). Near and above room temperature only the lesser slope F_{01} was observed and the conductivity and slope were found to be independent of thickness in the range[65] 10 μm to 1 mm and also in the range[64] of 0.2 to 1.1 μm.

The work in Ref. [56] investigated subthreshold d.c. conduction in modern GST PCM devices with thicknesses on the order of 100 nm and over a temperature range of 25 to 85 °C. Cell geometries included lance (vertical with 'hemispherical' amorphous dome) and μtrench[66] configurations. An ohmic regime was observed in the IV data at applied voltages $V_a < 0.3$ V (or fields $F \lesssim 3 \times 10^4$ V/cm), followed by an apparent

exponential dependence of $\ln I \propto V$. In the non-ohmic region, the activation energy was found to decrease linearly from 0.35 to 0.28 eV with increasing voltage. At yet higher voltages of 0.8 to 1 V, the slope of the $\ln I$ vs. V curves were found to be inversely proportional to temperature. These IV characteristics and temperature dependencies were corroborated in Ref. [55] for 100 nm thick amorphous GST films (and other compositions).

In contrast to the above results, a recent study[18] of GST PCM lance devices in the reset state, with thicknesses of less than 50 nm, presented IV data that was best described by $\ln I \propto \sqrt{V}$. An ohmic region was not observed but we note, however, that for such thin devices a field of 10^4 V/cm corresponds to an applied voltage of 0.05 V, below which data was not presented; hence, the low field region may not have been studied. Results of other work[50] for as-deposited amorphous GST films with thicknesses of 20 to 100 nm showed the same $\sqrt{V}$ dependence at fields of $F > 10^4$ V/cm and an ohmic regime for lower fields. The slope of the $\ln I$ vs. $\sqrt{V}$ curves increased slightly with temperature in the range 295 to 323 K (opposite to the above discussed results in Ref. [56]). The IV data in both Refs. [50] and [18] show a stronger field dependence near the threshold field.

A systematic investigation of thickness-dependent effects in thin glassy films was reported in Ref. [52]. Low-field ohmic and a high-field non-ohmic $\ln I$ vs. $\sqrt{F/F_0}$ regions were observed in IV data over a temperature range of 130 to 373 K and a thickness range of 130 to 600 nm for various compositions of GeBiSbS alloys. The measured current showed only a weak dependence on sample thickness and the slope, F_0, was thickness independent.

More recently, [67] a nonlinear relationship between low-field resistance and thickness was reported for μtrench GST PCM cells with amorphous GST thicknesses between 8 and 35 nm. However, the different thicknesses, which were obtained by varying the reset pulse, were calculated from the IV data by assuming the correctness

of a modified Poole-Frenkel conduction mechanism. The reported thickness dependencies do not account for the possible leakiness of such thin samples. That data is further discussed in Sec. 4.4 below.

In summary, as shown in Fig. 4-1, the experimental data suggest that there are three major field-dependent regimes: 1) an ohmic region at low field $F \lesssim 10^3 - 10^4$ V/cm (which can also be the start of a subsequent non-ohmic dependence); 2) an exponentially field dependent regime (possibly two distinct relations depending on temperature and thickness), with $\ln I \propto V$ or $\propto \sqrt{V}$ behavior, or some combination[68] thereof; and 3) a stronger field dependence, possibly $\ln I \propto V^2$, near the threshold field. The ohmic region may not be observed for very thin devices (e.g. thickness less than 100 nm) since the transition field is directly proportional to the thickness. There is evidence that the activation energy decreases linearly with increasing voltage in the intermediate field region. In general, the conductivity at or above room temperature appears to be nearly thickness independent, at least down to 100 nm.

4.2 Electronic States in Chalcogenide Glasses

In this section we recall the unique nature of charge transport in amorphous materials and the peculiar features of localized states in chalcogenide glasses. The electronic structure of disordered systems is inherently different than crystalline materials. The most important distinctive property of disordered materials is the short mean free path l that can be as small as the minimum quantum limit defined by $kl \gtrsim 1$, where k is the wavenumber. The strong scattering is due to intrinsic imperfections and random fluctuations of the potential energy.[35]

Since scattering in the disordered system is strong enough, $kl \sim 1$, the electronic wavefunction is effectively 'localized' to exist within a span of its wavelength. The celebrated work by Anderson in 1958 showed that random fluctuations in local site energies can create these localized states from which the wavefunctions fall off ex-

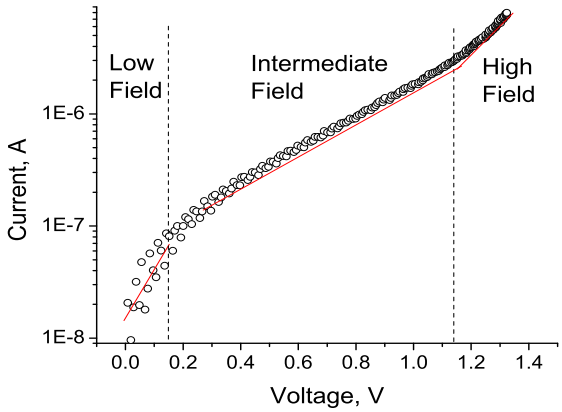

Figure 4-1: Three regions in the IV characteristic of an unspecified chalcogenide PCM device that is representative of the results discussed in the experimental overview. The low-field region is usually described as ohmic, but in some cases of thin samples it is described as $\ln I \propto V$. The intermediate region has exponential dependence described as either $\ln I \propto V$ and/or $\propto \sqrt{V}$. Near and below room temperature, two slopes are often observed in the intermediate region. The high field region corresponds to a stronger dependence, possibly $\ln I \propto V^2$.

ponentially with distance. Although these states share many similarities with traps in crystalline materials, the unique feature is that in amorphous materials there can be continuous distributions of such states that remain localized even if neighboring wavefunctions overlap. As a result, with sufficient disorder all diffusive transport can cease (except for thermal activation) and the conductivity will tend to zero at the zero of temperature even if localized states exist at the Fermi energy. The possibility of thermal activation facilitates hopping diffusion via localized states. The corresponding hopping conduction can be efficient enough to dominate over the band transport in such materials as a-Si and a-Ge. However it is significantly suppressed in chalcogenide glasses due to a very unique nature of their localized states as explained next.

Within a broader scope, deviations from structural periodicity smear out the sharp

energy bands into smooth transitions between localized and non-localized states that are separated by a distinct energy level, referred to as the mobility edge.[35] Furthermore, disorder and structural defects (such as dangling bonds) can lead to a continuous localized energy spectrum in the mobility gap and a finite density of states (DOS) at the Fermi energy (see Figs. 4-5 and 4-9). While that is true of any amorphous structure, the uniqueness of glassy semiconductors stems from the 'softness' of the atomic lattice, wherein the local atomic configuration can change significantly depending on the occupation number of the localized state. Hence, one must consider the consequences of electron-lattice interactions due to strong polaron effects, which, in particular, can significantly suppress the hopping transport. A summary of the observed phenomena and corresponding theoretical explanations that are relevant to our discussion of conduction mechanisms are provided next.

4.2.1 Conflicting Observations

Thorough reviews of experimental data related to electronic transitions in chalcogenide glasses can be found in Ref. [35], with more limited reviews given later in Refs. [71] and [72]. Here we summarize some of the results that are pertinent to charge transport and the significance of localized electronic states.

Beyond the specific observations cited in Sec. 4.1, experimental data on the various electronic properties of chalcogenide glasses can be broken into two groups, one of which testifies in favor of a high DOS in the mobility gap, while another states the opposite. To explain the data that suggests the high DOS, we consider the band diagrams shown in Fig. 4-2, which assume the standard one-electron localized states associated with all energy levels in the mobility gap. In the top left diagram, the localized states can provide efficient screening of an external electric field (shown as the tilted band edges) by redistributing the localized electrons in such a way as to form a screening dipole layer. The left bottom diagram illustrates another property

of this system: strong electron spin resonance (ESR) associated with the states in the vicinity of the Fermi level occupied by single electrons whose spins can be aligned with the external magnetic field (electron states well beyond the Fermi level can be occupied by pairs of electrons with opposite spins that do not contribute to ESR). The dashed arrows in the top right diagram show the transitions corresponding to a considerable (proportional to the high density of localized electron states) optical absorption for the photon energies $\hbar\omega$ smaller than the mobility gap G. Finally, the bottom right diagram illustrates hopping conduction via localized states close to the Fermi level. All the above phenomena - strong screening and ESR, noticeable absorption at $\hbar\omega < G$, and hopping - are observed in the tetrahedral amorphous semiconductors a-Si and a-Ge, for which the model of a high DOS of one-electron localized states in the mobility gap then appears fully adequate and comfortably self-consistent.

The conflicts arise when the above model is applied to chalcogenide glasses. It was observed that similar to a-Si, strong screening of the electrostatic field takes place, thus testifying in favor of a high DOS in the mobility gap (top left diagram in Fig. 4-2). On the other hand, the ESR signal is practically absent, thereby challenging the illustration in the left bottom diagram and shedding doubt on the presence of localized states in the mobility gap. However, a strong ESR signal can be induced by well-absorbed light; this photo-ESR is consistent with the hypothesis of a high DOS in the mobility gap. The optical transitions shown with dashed arrows in the top right diagram were not observed, suggesting that there may be no states in the gap. Instead, the transitions shown with solid arrows were observed, one of which corresponds to the interband absorption (not requiring localized states), while another one, downward, represents photoluminescence with energies around $\hbar\omega = G/2$. The latter implies a high concentration of localized states close to the Fermi level. Finally, as opposed to the case of a-Si, no hopping conduction was observed in chalcogenide

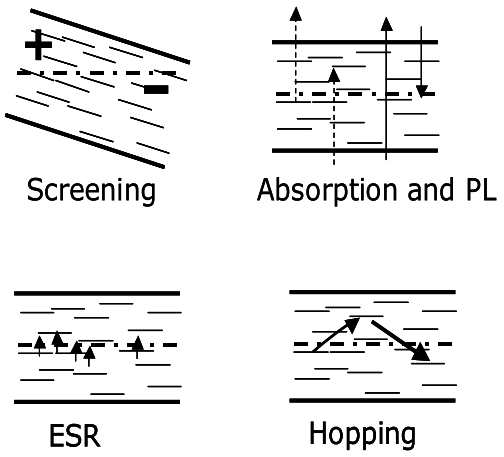

Figure 4-2: Sketches of physical processes associated with the one-electron localized states model. Top left: screening in the presence of an applied field due to redistribution of localized electrons to form a dipole. Bottom left: unpaired electrons near the Fermi level (dash-dot line) produce a strong ESR signal. Top right: optical absorption of photon energies less then (dashed arrows) and greater than (solid upward arrow) the gap, and photoluminescence (downward arrow) possible at mid-gap energies - solid arrows show what is observed in chalcogenide glasses. Bottom right: hopping conduction via states near the Fermi level. The one-electron localized states model cannot consistently account for the data on chalcogenide glasses.

glasses,[35] testifying against the model of a high DOS in the gap.

As a result, the group of observations against a high concentration of localized states includes: lack of ESR signal, absence of hopping conduction, optical gap G_0 approximately equal to the mobility gap G, and relatively low absorption of photons with energy less than G. On the other hand, the group in favor of a high concentration of localized states includes: strong photoinduced ESR corresponding to the electron concentration $\lesssim 10^{20}$ cm^{-3}, photoluminescence with energy close to $G/2$, d.c. screening length revealing a DOS at the Fermi level of $10^{18} - 10^{19}$ cm^{-3} eV^{-1}, strong pinning of the Fermi level close to the mid gap, photoinduced mid-gap absorption, and photoinduced change in the mid-gap photoluminescence. The spectroscopic aspects of these facts are illustrated in Fig. 4-3.

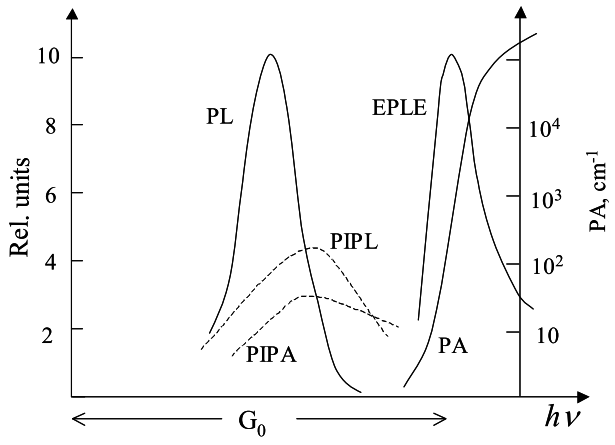

Figure 4-3: Sketch of the typical spectroscopic data in chalcogenide glasses: photoabsorption (PA), photoluminescence (PL), efficiency of photoluminescence excitation (EPLE), photoinduced PL enhancement (PIPL), and photoinduced photoabsorption (PIPA) vs. photon energy $h\nu$. G_0 is the optical gap. All the curves except PA are plotted against the left vertical axis.

4.2.2 The negative-U Model and Soft Atomic Potentials

A solution to the above controversy was proposed by Anderson [70] who put forward the concept of negative-U (negative Hubbard or negative correlation) energy which implies that two identical charge carriers localized at the same center will attract in spite of the Coulomb repulsion. As a result, double occupancy of a localized state becomes energetically more favorable than single occupancy of two localized states, such that the equilibrium occupation is $n = 2$ (electrons or holes), while $n = 1$ can only exist as an excited state. The conceptual leap of the negative-U model in chalcogenide glasses is that electronic-lattice interaction can be so strong that the energy is minimized when a localized state is *double occupied* and surrounded by a self-consistent cloud of lattice deformation.

Since the negative-U model favors two-electron states, it obviously explains the inadequacy of the one-electron model to account for the conflicting observations. For example, the lack of ESR in spite of a high concentration of localized states is due to

the fact that the states near the Fermi level are doubly occupied. The photoinduced effects become attributable to the non-equilibrium, single-occupancy states excited by higher energy photons. The rest of the above listed observations can be understood when the nature of the negative-U energy is specified as being related to an abnormally strong electron-lattice interaction for localized charged carriers.

The energy of $n = 0, 1, 2$ localized carriers is described as,

$$E_n(x) = nE_0 + kx^2/2 - nQx + U_c\delta_{n,2}, \tag{4.1}$$

where E_0 is the bare energy of the center, x is the lattice deformation around the center, k is the corresponding spring constant, Q is the deformation potential for the localized carriers, and U_c is the Coulomb repulsion energy applicable when $n = 2$. The dependencies in Eq. (4.1) are illustrated in Fig. 4-4. It should be understood that the bare energy levels E_0 are always present as unoccupied states near the mobility edge while states E_1 and E_2 are created by the lattice deformation that occurs when a bare energy level becomes occupied. Therefore, states E_1 and E_2 do not exist in and of themselves but, rather, they are modified versions of the E_0 state.

The equilibrium energies are given by the equation,

$$E_n = nE_0 - n^2w + U_c\delta_{n,2} \quad \text{with} \quad w \equiv \frac{Q^2}{2k}, \tag{4.2}$$

where w is called the polaron shift which quantifies the strength of electron-lattice interactions. From Eq. (4.2), the correlation energy is given by,

$$U \equiv E_2 - 2E_1 = -2w + U_c. \tag{4.3}$$

The postulated negative value of the correlation energy corresponds to a strong polaron effect with $w > U_c/2$.

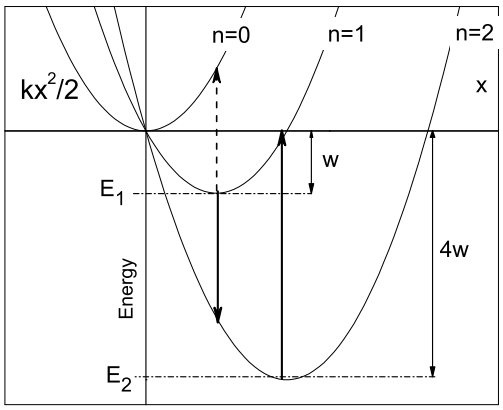

Figure 4-4: Energies of n localized charge carriers vs. the local lattice deformation x. The upward solid arrows represent absorption and the downward solid arrows represent photoluminescence processes; the dashed arrow indicates photoinduced photoabsorption from the nonequilibrium $n = 1$ state. E_1 and E_2 represent the equilibrium energies for $n = 1$ and $n = 2$ localized carriers. w is the polaron shift and U_c is assumed to be relatively small.

By the Franck-Condon principle, the characteristic energy of the absorbed light in Fig. 4-4 is $|E_2|$ while that of emission (PL) and photo-induced absorption is $2|E_1|$, and, assuming U_c relatively small, $|E_2| \approx 4|E_1|$, consistent with the data in Fig. 4-3. These transitions are shown in Fig. 4-5 with respect to the mobility gap. Note a significant Stokes shift (difference between the absorbed and emitted energies) approximately equal $G/2$ caused by the strong electron-lattice interaction. Comparing Figs. 4-3, 4-4, and 4-5 enables one to estimate $w \approx G/4$ (although w can be somewhat different for the cases of electrons and holes [35]). Also, note that thermal and optical transitions correspond to different positions of the energy levels in the mobility gap (see Fig. 4-5). The reason for this difference is that optical transitions occur at rates that are too rapid for lattice deformations to occur while thermal transitions allow sufficient time for the lattice to relax, resulting in a change of the energy level before the transition occurs. These differences illustrate the importance of electron-lattice

interactions and how the deformations affect the properties of glasses.

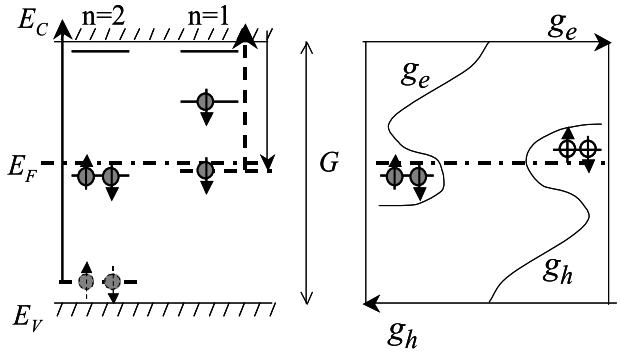

Figure 4-5: Left: one-particle energy levels (i.e. energy per particle) corresponding to $n = 2$ and $n = 1$ electrons in the mobility gap. The levels without electrons represent the bare energy. Solid and dashed lines indicate thermodynamic and optical energy levels, respectively. The dashed electron level close to the valence band edge represents the energy needed to optically ionize the 2e state (solid upward arrow); the solid level close to the midgap represents the energy needed to thermally ionize the same 2e state. The arrows have the same meaning as in Fig. 4-4. Right: density of the 2-electron (g_e) and 2-hole (g_h) states vs. their one-particle energies where negative-U centres near the Fermi level provide its pinning.

Drawing similar energy levels for holes and allowing for some dispersion, leads to the right diagram in Fig. 4-5 that explains how the Fermi level is pinned by a high concentration of 2e and 2h states, forming a gapless spectrum of two-particle excitations. In addition to the pinning (2e,2h) states, shown in the same diagram are band tails possessing the characteristic decay scales on the order of several hundredths of eV. They can contribute to optical absorption and act as shallow traps underlying dispersive transport and other phenomena. [71] As shown in Fig. 4-5, the one-particle excited states 1e and 1h are obtained through the partial ionization of (2e,2h). Possessing energies of approximately $w = G/4$ from the corresponding mobility edges, they can affect transport phenomena. [72]

4.2.3 The Nature of Negative-U Phenomenon

The microscopic nature of negative-U centers is not particularly important for the purposes of this work; here we limit ourselves to a brief comment on the subject. We note that the negative-U phenomenon can be simply illustrated in terms of a mechanical analogy with two electrically charged balls, each of weight Q, that can be attached to either two different elastic springs or one such spring, as depicted in Fig. 4-6. The spring elongations represent lattice deformations and the potential energy of the springs is related to the polaron shift w. The scenario with two charged balls on one spring turns out to be energetically favorable when $w > U_c/2$. A related generic interpretation of pairing in terms of the number of electrons occupying a dangling bond, as shown in the bottom row of Fig. 4-6, does not explicitly show the lattice deformation.

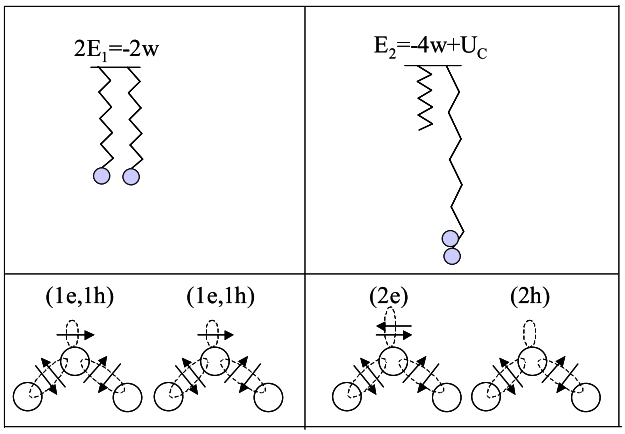

Figure 4-6: Mechanical analogy of the negative-U effect consisting of two elastic springs and two charged balls that can be attached to the springs either separately or together (top row) and its simple model based on the valence bonds representation (bottom row) where two electrons can occupy the states of two broken bonds or one dangling bond. The right column is energetically more favorable when $w > U_c/2$.

Street and Mott [73] proposed a microscopic model where 2e and 2h states correspond to certain defect states (D^- and D^+), while 1e and 1h is the same dangling

bond (D^0). Kastner *et. al* and Kastner and Fritzsche [74] introduced more specific consideration taking into account the chemical nature of chalcogenide forming atoms; in their popular notation D^- and D^+ are represented as C^{1-} and C^{3+} where the superscript indices refer to defect coordination numbers.

Later work [75, 76, 77, 71] emphasized that a theoretical description of the negative-U must explain the observed strong Stokes shift and, hence, the underlying significant polaron shift w. The required shift was attributed to centers with abnormally small spring constants k (soft atomic potentials) that exist in glasses due to their inherent structural disorder. Qualitatively speaking, the soft atomic potentials represent small pockets of a very soft liquid-like phase arrested during the vitrification process in the macroscopically solidified glassy structure. Because they are abnormally soft, these small inclusions interact abnormally strongly with the charge carriers giving rise to the abnormally large polaron shifts.

The random nature of a glass structure implies that the local spring constants are continuously distributed as illustrated in Fig. 4-7. Correspondingly, there exists a continuous distribution of local polaron shifts w including those responsible for the states in the proximity of the Fermi level. In particular, the polaron shift $w \approx G/4$ implies the spring constant $k_G \approx \langle k \rangle (2\langle w \rangle / G) \ll \langle k \rangle$ where $\langle w \rangle \lesssim 0.1$ eV is the average polaron shift corresponding to the average spring constant $\langle k \rangle$ which describes the macroscopic properties of glasses. We note that the same concept of soft atomic potentials has successfully explained the presence of atomic double well potentials (DWP) and localized quasiharmonic local vibrations in glasses. [76, 77, 78]

A comment is in order regarding the region of very small spring constants k that may seem to result in the infinitely large polaron shifts capable of localizing many ($n > 2$) charge carriers in the same microscopic region. It should be understood in this connection that for small enough k, the harmonic approximation for atomic potential fails and anharmonic terms become important. In fact, the soft atomic

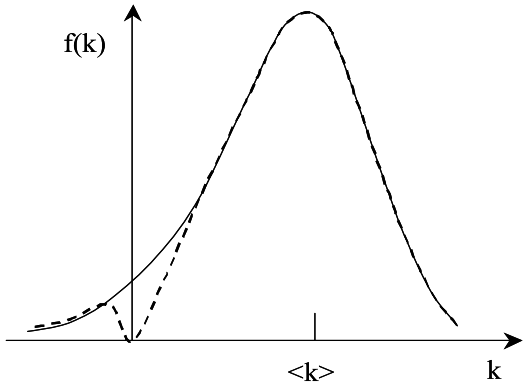

Figure 4-7: Probabilistic distribution of the local spring constants in a glass. The gull-wing singularity at the origin (not particularly important in this context) reflects the instability of very soft potentials with respect to small perturbations. [77]

potentials are described by the expression [75, 76, 77, 71]

$$V(x) = \frac{kx^2}{2} + Bx^3 + Cx^4 \tag{4.4}$$

where k and B are random quantities much smaller than their average values, while C is about its average (finding all the parameters noticeably off their respective average values would be extremely unlikely). It follows then that the harmonic approximation is limited to $k > k_h \equiv (2Q^2C)^{1/3}$. For lower k, the anharmonic term Cx^4 would govern the electron auto-localization, and for $k \ll k_h$, the term $kx^2/2$ in Eq. (4.1) should be replaced with Cx^4, which results in the maximum polaron shift $w_{max} \approx 0.5(Q^4/C)^{1/3}$ not strong enough to allow the many carrier localization.

An important conceptual difference between the 'defect models' [73, 74] and that of soft atomic potentials [75, 76, 77, 71] is that the latter concentrates on the gigantic polaron shift that underlies the negative-U phenomenon, while the 'defect models' explain negative-U by means of specific defect electronic orbitals, leaving the observed gigantic Stokes' shift as an additional (postulated) property beyond that explana-

tion. However, the existence of gigantic polaron shift already explains the negative-U phenomenon [see Eq. (4.3)], making the 'defect models' excessive.

Another important feature of the soft-atomic-potential approach is that it deals with a continuous distribution of spring constants k having a tail down to $k_G \ll \langle k \rangle$, as illustrated in Fig. 4-7. Correspondingly, it assumes a considerable concentration of localized states between the mobility edges and the Fermi level for which the polaron effect is greater than the average but not as strong as required by the condition $U < 0$. Such states can show up under significant illumination by localizing unpaired electrons and holes, which can explain the observed photo-induced ESR, as well as the features of PIPL and PIPA in Fig. 4-3.

4.2.4 Electronic Transitions with Negative-U Centers

It should be remembered that the 2e and 2h gapless excitations typically have extremely long relaxation times related to the necessity of carrying a heavy polaron cloud (i.ẽ. atomic deformation) in the course of electron transitions. Here we will describe such slow transitions in terms of electronic double well potential (DWP) with a transition barrier W_B related to the polaron shift. The two minima of such a DWP will correspond to the charge states (0,0) and (2e,2h) of two centers with energies close to the Fermi level. Another term for DWP is 'two-level-system' (TLS).

To estimate the barrier height W_B for the electronic DWP, we consider in more detail the process of transformation from the (0,0) to (2e,2h) state of the two centers. We start with the (0,0) state, taking 2 electrons from the valence band up to conduction band which requires the energy loss of 2G. Placing them and the remaining two holes at their respective bare levels (very close to the band edges; see Fig. 4-5) results in a rather insignificant energy gain which we neglect here. Finally, letting the lattice locally deform will deepen the (one-particle) energies at the centers by $\approx G/2$ and, when multiplied by the N=4 carriers involved, this leads to the energy gain of 2G.

Overall, the total energy change is zero, while the maximum energy increase in the course of transition was $W_B = 2G$, which we identify with the DWP barrier height.

The rather large barrier of $W_B = 2G$ exponentially reduces the electron transition rate since it is proportional to $\exp(-2G/kT)$. Physically, the latter exponential is attributable to the low probability for thermal fluctuations to create the strong lattice deformations (polaron cloud) that must accompany such electronic transitions (we do not discuss here the low temperature effects that are governed by the zero point vibration energies instead of kT).

Hence, although there exists a high density of localized states near the Fermi level, hopping between those states is extremely unlikely due to the above described deformation related transition barrier $W_B \approx 2G$ (first estimated by Phillips [79]). The barrier is high enough to fully suppress d.c. hopping conduction that could occur through electron hopping between centers separated by distances on the order of the average inter-center distance.[35] Assuming for specificity $2G \approx 1.6$ eV for the case of GST glasses and implementing the standard estimates[35] yields the multiplier $\exp(-2G/kT) \sim 10^{-27}$. This predicts hopping conduction many orders of magnitude below what is observed in the non-glassy semiconductors (such as a-Si).

From another perspective, we note that if hopping conduction did occur in chalcogenide glasses it would have to proceed through the above described excitations which occur within the mobility gap; via the negative-U centers near the Fermi level with their associated lattice deformations. On the other hand, band conduction is an interband mechanism which forgoes the restrictive lattice deformation process. Therefore, the conductivity for band conduction retains the thermally activated factor $\exp(-G/kT)$.

In the estimate of the transition barrier, we have neglected both the quantum contribution caused by the overlap of the wave functions of spatially close 2e and 2h centers and the Coulomb interaction of 2e and 2h pairs. It was shown [80] that

both corrections are significant for the case of *spatially close* pairs, sometimes called intimate pairs, which can decrease the barrier height by several times. In particular, the intimate (2e,2h) pairs partially decrease their energy due to the strong Coulomb interaction, which relaxes the requirement of very soft atomic potentials with $k \approx k_G$. Because higher k values result in smaller w, the factor $\exp(-2G/kT) \approx \exp(-8w/kT)$ describing the suppressing effect of a polaron cloud on the electronic transition becomes less significant, allowing for much higher hopping probabilities. However, we note that since the magnitudes of the latter and quantum effects depend on the details of the atomic and electronic structure of negative-U centers, our knowledge about W_B for such compact pairs remains rather approximate.

Lacking more accurate information, one can resort to the data [81] on alternating current (a.c.) conduction in chalcogenide glasses that is comparable to that of other noncrystalline semiconductors at relatively low frequencies ($\omega \gtrsim 1$ kHz). Because a.c. conduction is attributed to electron hopping between close centers,[81] these observations can be explained assuming that the electron transitions in *close pairs* of negative-U centers are as efficient as in the pairs of centers without a strong polaron effect, such as in a-Si. The assumed effectiveness of the electronic transitions in close pairs can be explained [80] by significant suppression of the transition polaron-related barrier in intimate pairs. Qualitatively, such a suppression is due to a substantial spatial overlap of the polaron clouds of two close centers, which makes it unnecessary to fully dissipate and recreate a polaron cloud in the course of electronic transition. The fact that no d.c. hopping was observed in chalcogenide glasses [35] means then that no barrier suppression takes place for the centers at distances close to the average.

The activation relaxation time for the electronic DWP formed by a pair of negative-U centers can be estimated as

$$\tau = \tau_{min} \exp\left[\frac{2R}{a} + \frac{\Delta W_B(R)}{kT}\right], \tag{4.5}$$

with

$$\tau_{min} = \tau_0 \exp\left[\frac{W_B(R_{min})}{kT}\right], \tag{4.6}$$

and

$$\Delta W_B = W_B(R) - W_B(R_{min}), \quad (\Delta W_B)_{max} \approx 2G \tag{4.7}$$

where R is the intercenter distance, a is the electron localization radius at the center, $\exp(2R/a)$ describes the electron tunneling, and $W_B(R_{min})$ is the activation barrier for intimate pairs separated by the distance $R_{min} \sim a$. Because R is a random quantity with the probabilistic distribution $4\pi R^2 N_U$ where N_U is the concentration of negative-U centers, the probabilistic distribution of relaxation times becomes,

$$\rho(E,\tau) = \frac{P}{\tau}, \quad P \approx \frac{\pi N_U^2 a^3}{4\Delta E_U} \frac{[\ln(\tau/\tau_{min})]^2}{1 + 2G/kT}, \tag{4.8}$$

where we have used a rough estimate $dW_B/dR \sim (\Delta W_B)_{max}/a$ and where ΔE_U is the total energy width of the negative-U center distributions that are approximately uniform in the proximity of the Fermi level. Neglecting the logarithmically weak dependence of P vs. τ, it can be treated as a constant.

For numerical estimates we use the values discussed in Section 9.4 of Ref. [35] that suggest $N_U \sim 10^{17} - 10^{18}$ cm^{-3} and $\Delta E_U \sim 0.025$ eV, yielding $P \sim 10^{15} - 10^{17}$ eV^{-1} cm^{-3}. The transition time τ corresponding to the typical $a \sim 10$ Å, average $R \sim N_U^{-1/3}$, and $W_B \approx 2G \sim 2$ eV turns out to be long enough ($\sim 10^{17} - 10^{25}$ s) to fully suppress hopping conduction. [79] Indeed, the latter can be estimated as $\sigma \sim e^2/(\tau kTR) \sim 10^{-26} - 10^{-34}$ Ω^{-1}cm^{-1}, much less than the experimentally observed $\sigma \gtrsim 10^{-3}$ Ω^{-1}cm^{-1}. We note that the frequency-dependent a.c. conduction at $\omega \gtrsim 1$ kHz is yet far enough from the limiting case of low frequencies bordering the d.c. regime: the latter would take place for $\omega \sim 1/\tau \sim 10^{-17} - 10^{-25}$ Hz. This range of incredibly small frequencies is due to the strong polaron effect that exponentially slows down the electronic transitions.

On the other hand, spatially close (intimate) pairs can have much lower W_B and exponentially shorter relaxation times than the average distant pairs, thus making noticeable contributions to the system noise[82] and a.c. transport in a broad range of relatively low frequencies. One other important property of intimate pairs is that they can form untypical chains between the electrodes of very thin samples (see Sec. 4.3.6). The distance between the centers in such chains will be much shorter than the average making them rather efficient channels for d.c. conduction. Therefore, one can expect that extremely small devices can exhibit transport properties significantly different from their larger counterparts. We shall see in what follows that such untypical transport can dominate conduction in the range of thickness well below 10 nm.

4.3 Survey of Conduction Mechanisms

The following subsections provide the physical basis, analytical expressions, and limiting assumptions for various non-ohmic conduction mechanisms. A summary of the expressions for the conductivity in each case is provided in Table 4.1.

Table 4.1: Listing of each conduction mechanism along with the related analytical expression and estimated field range of applicability. The current I is given in terms of the electric field F, with the pre-exponential $I_0 \propto (-E_a/kT)$.

Mechanism	$\ln(I/I_0)$	**Field Range (V/cm)**
Poole-Frenkel 1-center activation	$\frac{2}{kT}\sqrt{\frac{q^3F}{\varepsilon}}$	$10^4 - 10^5$
Continued...		

Table 4.1: (continued)

Mechanism	$\ln(I/I_0)$	**Field Range (V/cm)**
Poole-Frenkel 2-center activation	$\frac{aqF}{kT}$	$< 10^4$
Poole-Frenkel 1-center tunneling	$\frac{\hbar q^2 F^2}{3m}\left(\frac{1}{kT}+\frac{1}{kT_{ph}}\right)^2$	$> 10^5$
Schottky emission	$\frac{1}{kT}\sqrt{\frac{q^3F}{\varepsilon}}$	n/a
Delocalization of tail states	$\left(\frac{\hbar qF}{\sqrt{m}}\right)^{2/3}\left(\frac{1}{kT}-\frac{1}{E_0}\right)$	$\sim 10^5$
Space-charge limited currents	$\frac{\varepsilon F}{2\pi Lqgk T}$	$\sim 10^4$
Optimum channel hopping, thin films	$-\sqrt{\frac{8L\lambda}{\alpha}}+1.6\sqrt{\frac{qFL}{kT}}$	$< \frac{E_F}{qL}$
Optimum channel field emission	$-\sqrt{\frac{8\lambda E_F}{\alpha qF}}$	$\gg \frac{E_F}{qL}$
Percolation band conduction	$\eta\sqrt{\frac{L_c qF}{kT}}$	$> 10^4$
Percolation band conduction thin films ($L < L_c$)	$\eta\sqrt{\frac{L_c qF}{kT}}+\frac{L_c-L}{2r_c}\left[\ln\left(\frac{2V_{max}r_c}{kTL}\right)+1\right]$	$> 10^4$
Crystalline inclusions (1)	$\frac{2\varepsilon}{kT}\sqrt{\left(\frac{r_x\Delta}{q}\right)^3 F}$	$10^5 - 10^6$
Continued...		

Table 4.1: (continued)

Mechanism	$\ln(I/I_0)$	**Field Range (V/cm)**
Crystalline inclusions (2)	$\frac{\varepsilon r_x^2 F \Delta}{qkT}$	$< 10^5$

The parameters are defined as follows: k is the Boltzmann constant, T is temperature, q is the elementary charge, ε is the dielectric constant, a is the inter-center distance, $\hbar$ is the reduced Planck's constant, m is the effective carrier mass, $kT_{ph} \sim 0.01 - 0.03$ eV is the characteristic phonon energy, E_0 is the characteristic decay of the density of tail states $g = g_0 \exp(-E/E_0)$, where E is energy, L is thickness, $\lambda \approx -\ln(g_0 kTaL^2) \gg 1$ (here, g_0 is the density of localized states), α is the electron localization radius, E_F is the Fermi energy, $\eta \sim 1$ is a numerical factor, $L_c \sim 10$ nm is the percolation cluster correlation radius, r_c is the order parameter, r_x is the crystallite radius, V_{max} is the maximum percolation transport barrier, and $\Delta \sim 0.4$ eV is the band offset between crystalline and amorphous phases.

4.3.1 Poole-Frenkel Effect

The originally suggested physics of the PF effect is the decrease in the ionization energy of a single coulombic potential well in the direction of an applied field (explaining $\ln I \propto \sqrt{V}$) or that of a pair of coulombic centers (explaining $\ln I \propto V$), as illustrated in Fig. 4-8. The corresponding barrier change δ increases the center ionization rate, proportional to which are the free carrier concentration and the activated electric current $I/I_0 \propto \exp(\delta/kT)$. The underlying assumption of a coulombic attractive potential is justified by its ability to give the required decrease in the ionization

energy $\delta \propto F$ or $\delta \propto \sqrt{F}$.

We note that as originally proposed, [48] this mechanism was meant to explain the data on noncrystalline materials (mica, SiO_x, etc.; see Ref. [47] and references therein). Surprisingly, the data on non-ohmic conduction in doped crystalline semiconductors are typically described by other dependencies, [83] despite the fact that the coulomb nature of the defects therein is well established. Therefore, the empirically observed relevance of PF-type dependencies to noncrystalline materials may suggest that their nature is more related to disorder effects rather than individual or pairs of coulomb centers. From that point of view, the PF mechanism may be significantly over-emphasized.

For the case of two centers separated by distance $2a$ in the electric field of strength F, the electron energy along the axis is given by,

$$U(x) = -\frac{q^2}{\varepsilon(a-x)} - \frac{q^2}{\varepsilon(a+x)} - Fqx, \tag{4.9}$$

where q is the electron charge and ε is the dielectric permittivity. The position of the lowest barrier maximum, $dU/dx = 0$ is determined from the equation

$$\tilde{x} = \tilde{F}(\tilde{x}^2 - 1)^2 \quad \text{where} \quad \tilde{x} = \frac{x}{a}, \quad \tilde{F} = \frac{F}{q/4a^2\varepsilon}. \tag{4.10}$$

The original PF result $x = \sqrt{q/\varepsilon F}, \quad \delta = \sqrt{4q^3F/\varepsilon}$ follows from Eqs. (4.9) and (4.10) when $\tilde{F} \gg 1$ (i.e. $F \gg q/4\varepsilon a^2$); however, it remains approximately valid numerically even at $\tilde{F} = 1$. The characteristic field is $q/4a^2\varepsilon \sim 10^4$ V/cm for the typically assumed [35, 56] center concentration of $\sim 10^{18}$ cm^{-3}.

In the opposite limiting case of 'weak' fields, $F \ll q/4\varepsilon a^2$, Eqs. (4.9) and (4.10) yield $x = a$ and $\delta = qFa$, corresponding to the so called modified PF effect with $\ln(I/I_0) \propto V$ emphasized in Ref. [56].

The critical field $q/4\varepsilon a^2$ also implies that the notion of 'weak' or 'strong' fields

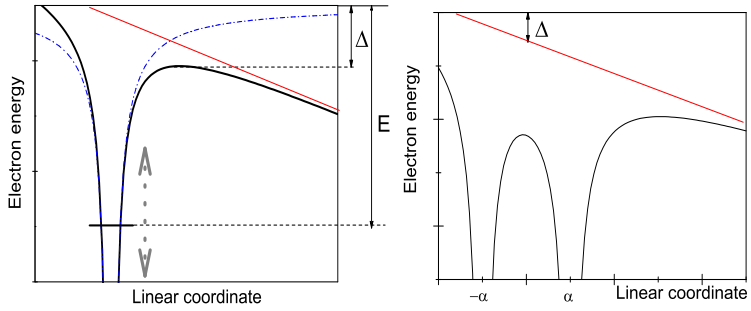

Figure 4-8: Left: Field induced decrease δ in activation energy of a coulombic center. Dashed lines show zero field case, tilted red line represents the electric potential of a uniform field. Gray arrow shows vibration of the electron energy E due to electron-phonon coupling. Right: Field induced decrease δ in activation energy of a pair of coulombic centers.

can be replaced by condition of low or high defect density (as related to a). Thus, for a given field, the PF effect is dominant for a high defect concentration while the modified PF effect pertains to a low defect concentration. In any event, we observe that significant deviations from the standard PF results can be expected under low fields $F \ll 10^4$ V/cm. This significantly narrows the application of the modified PF mechanism in Refs. [56] and [84], also ruling out its role in the switching field region of $F \gtrsim 10^5$ V/cm.

The two-center model that predicts $\ln(I/I_0) \propto V$ remains critically vulnerable to effects of fluctuations. We note in this connection that the work in Refs. [56] and [84] was limited to a system of equidistant coulombic centers. Random fluctuations in their concentration (present in all systems of centers in solids so far explored) will generate random variations of activation energies translating into exponentially broad distributions of ionization rates; variations in center energies will make this distribution even broader. This results in local carrier concentrations that vary exponentially

between different locations. A proper framework for analyzing these types of systems would be percolation theory, [85] which is yet to be applied to PF-type conduction (cf. however Ref. [86]).

Quantum tunneling imposes limitations on the activation PF effect. The corresponding analysis by Hill [47] neglects the role of atomic vibrations on tunneling. A more recent analysis [83] that accounts for electron-phonon interactions results in a picture where the electron energy level moves up and down following oscillations of the atomic system to which it is coupled. As a result, the electron tunneling becomes most likely when the electron energy is significantly above its average position (Fig. 4-8), and the chief exponential term in non-ohmic current is given by,

$$\ln(I/I_0) = \frac{F^2 q^2 \hbar}{3(kT^*)^2 m} \quad \text{with} \quad \frac{1}{kT^*} = \frac{1}{kT} + \frac{1}{kT_{ph}}, \tag{4.11}$$

where m is the effective mass of a localized charge carrier, which we take to be close to the true electron mass, [87] and kT_{ph} is on the order of the characteristic phonon energy ($\sim 0.01 - 0.03$ eV).

It was shown [83] that the standard PF results becomes invalid and the effect is better described by Eq. (4.11) when

$$F > F_t \equiv \sqrt{\frac{2mE}{\hbar^2}} \frac{kT^*}{q} \left(\frac{kT^*}{E}\right)^{1/3}, \tag{4.12}$$

where E is the ionization energy (≈ 0.4 eV in $Ge_2Sb_2Te_5$). Using the above numerical parameters, one can estimate $F_t \sim 10^5$ V/cm. We note that the dependence in Eq. (4.11), rather than the standard PF law, was experimentally confirmed for many crystalline semiconductors even for fields below 10^5 V/cm (see Chapter 10 in Ref. [83]).

Overall, we conclude that, for the case of GST glasses, the standard PF expression $\ln(I/I_0) \propto \sqrt{F}$ can apply in the field range of $10^4 - 10^5$ V/cm. For weak fields,

$F \ll 10^4$ V/cm, the modified PF effect $\ln(I/I_0) \propto F$ can give a more adequate description, however, the effects of fluctuations in the local concentration of centers must be taken into account. For the high field region, $F > 10^5$ V/cm, quantum effects lead to $\ln(I/I_0) \propto F^2$, predicting an increase in non-ohmicity in the vicinity of the switching field. The above boundaries can be numerically different for other chalcogenide glasses; however, the hierarchy of regimes remains the same, as illustrated in Fig. 4-1.

Experimental data[50, 18, 65] has exhibited a sharp increase in current when the field is very close to its switching value 3×10^5 V/cm, however, it would be premature at this stage to attribute it to Eq. (4.11). Experimental verification of the temperature dependence in Eq. (4.11) could clarify this issue.

4.3.2 Schottky Emission

The Schottky effect [88] originates from the image force induced lowering of the interfacial energy for charge carrier emission when an electric field is applied. This leads to,

$$\ln(I/I_0) = \frac{1}{kT}\sqrt{\frac{q^3F}{\varepsilon}} \quad \text{with} \quad I_0 \propto \exp(-\Phi/kT), \tag{4.13}$$

where Φ is the interfacial barrier height between the semiconductor and the contact metal.

The dependence in Eq. (4.13) was experimentally verified in the field range $\sim 10^4 - 10^5$ V/cm for various junctions of crystalline semiconductors with metals. However, on empirical grounds, it is hard to believe that it can apply to the case under consideration because of the established $\ln I_0 \propto (-E_a/kT)$, where E_a is half the mobility gap in the chalcogenide material and is independent of contact properties. Some studies reveal that the current is independent of polarity and electrode material, which is additional evidence against the Schottky mechanism. [52]

4.3.3 Field-Induced Delocalization of Tail States

Similar to the PF mechanism of decreasing the ionization energies of coulombic centers, the electric field can decrease energies of localized tail states in the mobility gap and even destroy them if they are shallow enough. Transforming localized into delocalized states is tantamount to narrowing the mobility gap; this exponentially increases the free carrier concentration and electric conductivity.

The latter mechanism, suggested in Ref. [58], is specific to noncrystalline materials where the presence of band tails is well established. Tail states are related to intrinsic structural disorder of amorphous materials rather than to any specific defects. The disorder creates microscopic variations in the electric potential generated by different structural units in a material and felt by electrons or holes. Some combinations of these microscopic variations form effective potential wells capable of localizing charge carriers.

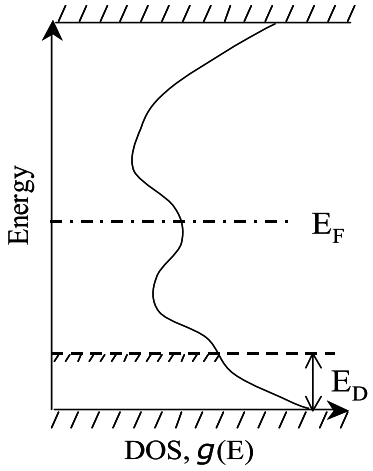

Figure 4-9: Density of states (DOS) in the mobility gap of a chalcogenide glass. The electric field shifts the mobility edge for holes up by energy E_D (similar effect for electrons is not shown here).

It was assumed in Ref. [58] that each fluctuation potential well has the same radius

r_0 regardless of the energy of its localized state, thus governed only by the well depth. Correspondingly, the condition of the electric field induced delocalization was given in the form $E < E_D \equiv Fqr_0$. Assuming also a simple phenomenological representation of the density of tail states, $g(E) = g_0 \exp(-E/E_0)$, the field-induced increase in concentration of charge carriers becomes $n(F) \propto g(E_D) \exp(E_D/kT)$, where the first multiplier describes the decrease in activation energy by E_D, as illustrated in Fig. 4-9. As a result, the conductivity increases with field as,

$$\sigma(F) = \sigma_0 \exp\left[Fqr_0\left(\frac{1}{kT} - \frac{1}{E_0}\right)\right], \tag{4.14}$$

where it is assumed that $E_0 > kT$. The observed temperature dependence in Ref. [58] was consistent with that in Eq. (4.14).

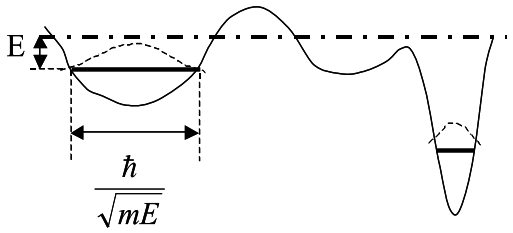

Figure 4-10: Localized tail states for the electrons below the mobility edge (shown as dash-dot line) have linear dimensions decreasing with energy E in the mobility gap.

The above model could be refined by taking into account that the characteristic size of the localized state of energy E is $\hbar/\sqrt{mE}$ and so is that of its corresponding potential well, [89] as illustrated in Fig. 4-10. As a result the condition of delocalization, approximately $Fq\hbar/\sqrt{mE} = E$, gives the characteristic delocalization energy $E_D = (\hbar qF/\sqrt{m})^{2/3}$ and, similar to Eq. (4.14),

$$\sigma(F) = \sigma_0 \exp\left[\left(\frac{\hbar qF}{\sqrt{m}}\right)^{2/3}\left(\frac{1}{kT} - \frac{1}{E_0}\right)\right]. \tag{4.15}$$

This prediction is in a numerically relevant range yielding $E_D \sim 0.1$ eV when $F \sim 10^5$ V/cm.

Further implementations of the theory of disordered systems [89] calls upon using the density of tail states in the form,

$$g(E) = g_0 \exp\left[-\left(\frac{E}{E_0}\right)^{\alpha}\right], \tag{4.16}$$

where $\alpha = 1/2$ and $\alpha = 2$ for the cases of uncorrelated and strongly correlated disorder corresponding respectively to the energies $E \ll \hbar^2/mr_c^2$ and $E \gg \hbar^2/mr_c^2$. If the correlation radius r_c is identified with that of the medium range order in a glass, [90] then $r_c \sim 1$ nm and $\hbar^2/mr_c^2 \sim 0.1$ eV. Using Eq. (4.16) will obviously modify the results in Eqs. (4.14) and (4.15) without changing them qualitatively.

Overall, it may be very difficult – if possible at all (see Sec. 4.4) – to experimentally discriminate between the shapes predicted by Eqs. (4.14) and (4.15) or their modifications. What is important is that these predictions pertain to a numerically relevant range $E_D \sim 0.1$ eV when $F \sim 10^5$ V/cm, ensuring strong enough non-ohmicity to explain the observed effects. Also, this model, in contrast to the PF model, gives a natural explanation of why PF-type non-ohmicity [$\ln(I/I_0) \propto \sqrt{F}$ or F] is typically observed in glasses rather than in crystalline materials.

4.3.4 Space Charge Limited Current

The exponential current-voltage characteristic can be explained by space charge limited current in a system with almost energy independent density of states.[91] This model is represented in Fig. 4-11 in the coordinate and energy spaces. Due to low mobility, the charge carriers accumulate in a system (the logarithm of their density is shown in Fig. 4-11 as the quasi-Fermi level) and create the potential barrier further slowing down their transport. In energy space, charge carriers occupy a layer of certain width δE near the Fermi energy (E_F). Therefore, their charge density is estimated

as $\rho = g(E_F)q\delta E$. The corresponding electrostatic potential is $V \approx 2\pi\rho L^2/\varepsilon$ where L is the sample thickness. Expressing from here δE through V and taking into account that the activation energy of conduction is by δE lower than in the ohmic regime, one gets

$$\sigma = \sigma(0)\exp\left(\frac{F}{F_0}\right) \quad \text{with} \quad F_0 = \frac{2\pi gqLkT}{\varepsilon}. \tag{4.17}$$

Assuming realistic $g = 10^{17}$ cm^{-3}eV^{-1} and $L = 100$ nm yields a relevant field scale of the non-ohmicity $F_0 \sim 10^4$ V/cm; however that scale strongly depends on the system thickness and density of states, which can make F_0 too large and irrelevant to the observed non-ohmicity in some chalcogenide glasses. The explanation of space charge limited current was put forward in Ref. [60] where F_0 linear in L was observed below room temperature. Near and above room temperature, F_0 was found to be thickness independent. [60, 64, 65] This data may suggest that space charge limited transport mechanisms play an important role in thicker samples ($L > 1$ μm) below room temperature.

We note that additional verification of the space charge limited mechanism of room temperature conduction in chalcogenide glasses could be obtained from the data on $1/f$ noise measurements. Our results [92] show that the corresponding Hooge parameter increases with bias, contrary to what is expected for the space charge limited currents. [93]

4.3.5 Hopping Conduction

The intent of this section is not to provide a complete description of hopping conduction, since thorough reviews are available elsewhere.[35] Here, we provide a brief explanation as to why hopping conduction was not observed experimentally in chalcogenide glasses.[35, 94]

A high density of localized states [g_F] at the Fermi level (E_F) in non-crystalline semiconductors can give rise to hopping transport. The mechanism is based on elec-

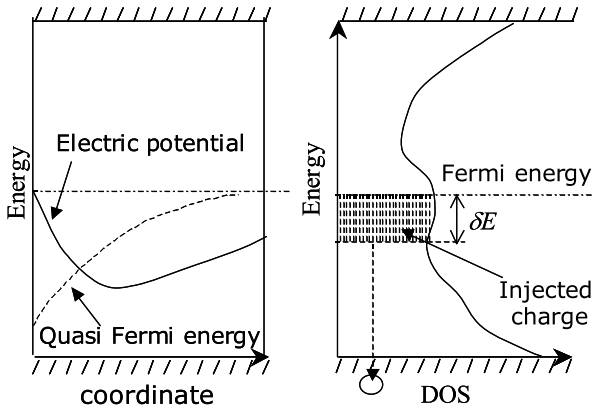

Figure 4-11: Left: real space representation of space charge (exponential in quasi-Fermi energy) and electric potential where the barrier top plays the role of a virtual cathode. Right: energy space representation with shaded region filled with injected holes.

tronic tunneling ('hops') between localized states that are randomly distributed in real space and energy space.[35, 58] In materials where hopping does occur, it dominates at low temperatures (T) and is described by the Mott law,[35]

$$\sigma = \sigma_0 \exp\left[-(T_0/T)^{1/4}\right], \quad T_0 = \beta/kg_F\alpha^3 \tag{4.18}$$

where α is the localization radius of the electron wave function, and $\beta \sim 1$ is a numerical factor. However, at room or higher T of practical interest, the primary transport mechanism in bulk materials is typically band conduction.

It has long been established that room temperature conduction in chalcogenide glasses is dominated by band transport.[35] One piece of evidence is that in all chalcogenide glasses the activation energy of conduction is close to half the mobility gap, $E_a \approx G/2$, identified with the Fermi level pinned at that position. As discussed in Sec. 4.2, lack of hopping is explained by the abnormally strong polaron effect for localized charge carriers [35, 70, 95] requiring electron transitions to be accompanied by

the inter-center transfer of atomic deformations (polaron cloud), which exponentially suppresses the probability of hopping. The strong polaron effect makes chalcogenide glasses significantly different from other amorphous semiconductors, such as a-Si, where hopping conduction was experimentally observed. [35, 96]

On a more quantitative level, we note that the polaron effect on hopping conduction was explicitly taken into account in Ref. [98]. It was shown [in Eq. (24) of that work] that in the high temperature regime the exponent of conductivity contains both the well known Mott term [35] $(T_0/T)^{1/4}$ and the polaron related term $W/2kT$ with the polaron shift W being close to $G/4$ as explained above in Sec. 4.2. The latter combination cannot be reduced to the observed activation conductivity exponent $\approx G/2$.

Finally, we note a simple estimate showing how hopping cannot provide the high current densities $j \sim 10^4$ A/cm^2 observed in the glassy state of modern PCM:

$$j \sim \frac{q\nu}{R^2} \exp\left(\frac{-E_a}{kT}\right) \sim 5 \quad \text{A/cm}^2, \tag{4.19}$$

where we have assumed the typical frequency of attempts $\nu \sim 10^{13}$ s^{-1}, inter-center distance $R \sim 10$ nm, and E_a =0.4 eV. For comparison, the devices of area 10^{-10} cm^2 with average current of 1 μA used in Ref. [56], correspond to a current density of 10^4 A/cm^2, decades higher than expected for hopping from Eq. (4.19).

The latter estimate can be put in a more standard perspective using Mott's criterion of band conduction, [35] according to which the thermally activated conduction $\sigma = \sigma_0 \exp(-E/kT)$ should have a preexponential in the range $\sigma_0 = 150 - 600$ Ω^{-1}cm^{-1}. Our data in Fig. 4-12 show that the latter criterion is satisfied for the case of GST based PCM.

Contrary to the above understanding, the authors of Refs. [56] and [84] proposed that conductivity in chalcogenide glasses is due to an altered form of hopping. In that work it was assumed that electrons move without tunneling between equally spaced

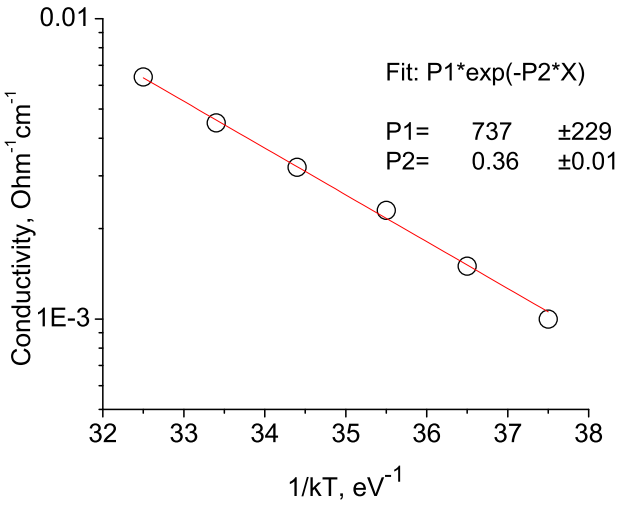

Figure 4-12: Temperature dependence of conductivity in a GST based PCM structure.

centers. The same hopping-without-tunneling mechanism was originally proposed for ionic conduction, i.e. for heavy (atomic) classical particles that possess continuous energy spectrum above the barrier. [33, 97] For the case of light quantum particles, such as electrons or holes, the spectrum is discrete and may have no quantum states between the barrier and the mobility edge.

The continuous energy spectrum needed for the purely activated transitions assumed in Refs. [56] and [84] starts at the mobility edge. Therefore, the 'no-tunneling' activated electronic transitions between the nearest neighbors would have to go via intermediate states at the mobility edge. However, allowing the electron or hole to utilize the states at the mobility edge is inconsistent with hopping conduction. Indeed, carriers at the mobility edge would attain the band mobility, which is well above that of hopping, thus giving rise to band transport and the nearest neighbor concept would not apply. In other words, having activated to the mobility edge, the charge carrier would become free and capable of traveling considerable distances to other (far from the nearest) traps or even to the device terminals. The above reasoning

explains why the hopping-without-tunneling mechanism has never been included in the existing theory of hopping conduction in semiconductors.

The work in Refs. [56] and [84] interpreted $E_a \approx G/2$ as the activation energy of hopping to the nearest center, assuming a transition through an intermediate state. In addition, it was assumed that all the inter-center distances are the same, thereby neglecting fluctuations in center concentration and activation energy, which are known to have exponentially strong effects on hopping conduction and determine the temperature and field dependence.[85]

4.3.6 Optimum Channel Hopping

Optimum channel hopping describes the gigantic transverse conduction that has been observed[101] in thin amorphous films. A thorough review of the related work is provided in Ref. [99]. Similar to classical hopping conduction discussed in Sec. 4.3.5, optimum channel hopping involves tunneling between localized states but it differs from the classical mechanism in the following ways: 1) optimum channel hopping does not occur on the macroscopically isotropic percolation cluster but, rather, through untypical and nearly rectilinear hopping chains of spatially close localized states; 2) it is characterized by laterally nonuniform (or pinhole) current flow; and 3) it can dominate over typical band transport in systems that are thin enough or subject to sufficiently strong electric fields. For chalcogenides, we consider the possibility that optimum channels can be comprised of localized states that are not subject to strong polaron effects.

Following the approaches in Refs. [99] and [101] we concentrate on optimum channel hopping through short distances via favorable yet sparse clusters of rather rigid localized states that form efficient transport pathways (see Fig. 4-13). The conductivity will be dominated by *optimum pathways* that are a compromise between a high transmission rate and not too low probability of finding the pathways being

considered. For the case of thin amorphous films, it was shown [99, 100] that optimum channel hopping leads to a transverse conductivity given by,

$$\sigma \approx \sigma_0 \exp\left(-\sqrt{\frac{8L\lambda}{\alpha}}\right), \tag{4.20}$$

where L is the thickness, α is the localization radius, $\lambda \approx -\ln(g_0 kT\alpha L^2) \gg 1$, and g_0 is the density of localized states.

Because Eq. (4.20) is not widely known to the microelectronic community, we mention here its simplified derivation. Consider a hopping pathway formed by N-center chain of almost equidistant centers. The probability of finding such a chain is estimated as $p^N = \exp(-\lambda N)$, where $\lambda \equiv \ln(1/p)$ and p is the probability of finding one center in the pathway. The probability of hopping through a distance L/N (between two nearest centers in the chain) can be written in the form $\exp(-2L/N\alpha)$, where α is the localization radius on the center. The product of these probabilities $\exp(-N\lambda - 2L/N\alpha)$ gives a partial current through an N-center chain. Optimizing it with respect to N determines the most efficient chains and results in Eq. (4.20); expressing λ through the density of states takes a more accurate approach. [99]

4.3.6.1 Optimum channels in thin films

For the case of thin amorphous films subject to moderate fields ($F < E_F/qL$, where E_F is the Fermi level), it was shown [99, 100] that optimum channel hopping leads to a transverse conductivity given by,

$$\sigma \approx \sigma_0 \exp\left(-\sqrt{\frac{8L\lambda}{\alpha}} + 1.6\sqrt{\frac{qFL}{kT}}\right), \tag{4.21}$$

where the parameters are the same as in Eq. (4.20). Polaron effects are neglected in Eq. (4.21) and, therefore, in chalcogenide glasses this form of hopping conduction cannot rely on the typical electronic states near the Fermi level. However, these

channels through extremely thin films or in the presence of strong fields (described more in detail in the next section) can be formed by untypical spatially close states, for which the effects of polaron cloud are less significant, or they can be formed by states far from the Fermi level having much smaller polaron shifts as explained in Sec. 4.2); for the case of chemically imperfect thin films, hopping in optimum channels could be due to extraneous states formed by certain impurities.

4.3.6.2 Optimum channel field emission

The standard interpretation of field emission is based on the model of electron tunneling through a triangular potential barrier with a slope F due to an electric field. [88] Our model here proceeds from the premise of a continuous energy spectrum of localized states in the mobility gap, typical of amorphous materials and capable of giving rise to hopping conduction. Such states lie high enough above the Fermi level that, according to the understanding in Sec. 4.2, they are not related to soft atomic potentials and thus do not posses the strong polaron shift that suppresses hopping. The possibility of hopping transport through such 'rigid' states far from the Fermi level is fully compatible with the above-described suppressed hopping at the Fermi level.

For the case [100] of strong fields, $F \gg E_F/qL$, Eq. (4.20) remains valid with the substitution $L \to l = E_F/qF$ (see Fig. 4-13 right). As a result, one obtains,

$$\sigma \approx \sigma_0 \exp\left(-\sqrt{\frac{8E_F\lambda}{\alpha qF}}\right), \tag{4.22}$$

which is significantly different from the standard field emission conduction with $\ln(\sigma/\sigma_0) \propto -1/F$.

One qualitatively distinctive feature of the above considered field emission is that it is significantly nonuniform and occurs through rare optimum channels (as opposed to the standard uniform Fowler-Nordheim emission from contacts [88, 102]); this may

lead to local heating, facilitating structural transformations in chalcogenide glasses. Another feature related to such lateral nonuniformity is that very small area devices, $A \lesssim \alpha L \exp(\sqrt{E_F \lambda / \alpha q F})$ may not have an optimum channel with certainty, in which case their resistances will be determined by the most efficient of available random channels; hence, there will be strong variations between the conductances of nominally identical cells.

Overall, it should be noted that the field emission mechanism can be expected to show up in very thin structures where the hopping resistance corresponding to Eq. (4.22) is not blocked by a significantly larger resistance of the film in series.

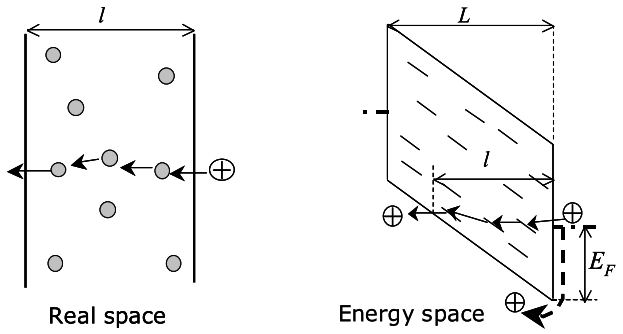

Figure 4-13: Left: field emission via hopping through an optimum chain; circles represent localized states. Right: same in the energy space.

4.3.7 Percolation Conduction

In general, conductivity in randomly nonuniform materials is described in terms of percolation. [85] This concept includes both the hopping conduction and band conduction in a medium where charge carrier concentration exponentially varies between different locations due to spatial variations in the electron potential energy. The concept of spatially varying mobility edge can be derived based on the above mentioned picture (Sec. 4.3.3) where a glass band structure is represented by a set of random

potential wells with localization/delocalization effects leading to the mobility edges. Some regions will contain predominantly deeper than the average or shallower than the average potential wells corresponding to local variations in the envelope electronic potential in the form of smooth wells or barriers. The latter variations translate into the electric conductivity exponentially varying in space.

Percolation conduction evolves on a mesh built of material regions with conductivity below a certain critical value $\sigma_c \equiv \sigma_0 \exp(-\xi_c)$ such that the mesh enables a connection between two flat electrodes, regardless of distance L between them. Such a mesh is called an infinite percolation cluster and is characterized by the correlation (mesh) radius $L_c < L$, as shown in Fig. 4-14.

The topology of the percolation cluster can be pictured as arising from a multitude of sites where the nearest neighbors can be connected with random resistors $R = R_0 \exp(\xi)$. Here ξ is a random parameter. For example, $\xi = E_F/kT$ for the case of band percolation conduction, where E_F represent a random energy distance between the band edge (which is spatially modulated) and the Fermi level. As another example, $\xi = 2\Delta r/\alpha + \Delta E/kT$ for hopping conduction, where Δr and ΔE are the distances between the two centers in the real and energy space respectively (α being the localization radius on the center). The cluster forming connection proceeds in sequence starting from the minimum resistor ($\xi = 0$) and adding larger ones up to $\xi = \xi_c$, until the everywhere connected cluster is formed. The mesh structure illustrated in Fig. 4-14 is built of filaments obtained by the series connection of random resistors where the maximum resistor is close to $R = R_0 \exp(\xi_c)$ for each of the filaments.

Following a theory of high-field percolation conduction,[103] each cell of the percolation cluster accommodates voltage $V_c = VL_c/L$. Because the resistors that constitute the filament are exponentially different, the latter voltage almost entirely concentrates on the strongest, first maximum resistor (1-max in Fig. 4-14). That voltage,

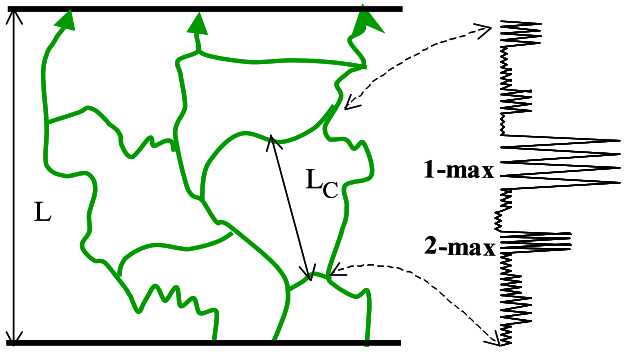

Figure 4-14: Left: Fragment of percolation cluster with mesh size L_c in a material of thickness L. Right: equivalent circuit of a filament of the percolation cluster where exponentially different resistors in series are depicted by resistors of different sizes; the first and second maximum resistors are marked for illustration.

concentrated locally, affects the resistance of the element across which it drops. The mechanism of the latter action can depend on the specific system under consideration: changes in center occupation numbers for the case of hopping, or field-induced ionization for the case of band transport. The field affected maximum resistor in the filament decreases its resistance down to the second maximum (2-max in Fig. 4-14), after which the voltage distributes evenly between the two resistors (1-max and 2-max), modifying both of them, and then extending to the third maximum resistor, etc. Such equalization will sequentially take place in a number of resistors having ξ_i from the maximum one (ξ_c) down to $\xi_0(V)$ defined by the condition,

$$\sum_{\xi_0}^{\xi_c} \xi_i = \frac{qV_c}{kT}.$$

Approximating the sum by the integral gives $(\xi_c - \xi_0)^2/2\xi_{max} = qV_c/kT$, where it is assumed that the random parameter ξ is uniformly distributed in the interval from 0

to $\xi_{max} \sim \xi_c$. As a result, the effective conduction is described by

$$\sigma \propto \exp(-\xi_0) = \exp\left(-\xi_c + \sqrt{\frac{2\xi_{max} q V_c}{kT}}\right).$$

Substituting here the definition $V_c = V L_c/L$ and $F = V/L$ one finally obtains

$$\sigma(F) = \sigma(0) \exp\left(\eta \sqrt{\frac{qFL_c}{kT}}\right), \tag{4.23}$$

where $\eta \sim 1$ is a numerical coefficient.

We observe that the conductivity depends on electric field in a manner very similar to the original PF result. Furthermore, assuming that each resistor has a linear dimension of the medium range order parameter r_c, L_c can be numerically estimated as $r_c(\delta E_a/kT) \sim 10 r_c \sim 10$ nm, where δE_a is interpreted as the amplitude of variations of the activation energy of conduction. It is estimated as the valence band offset between the most conductive (close to crystalline GST) and least conductive amorphous GST regions: $\delta E_a \approx 0.4$ eV. With the above estimate in mind, Eq. (4.23) predicts significant non-ohmicity starting from $F \sim 3 \times 10^4$ V/cm, in reasonable agreement with observations.

Finally, we note that in the case of very thin films, $L < L_c$, the transversal conduction will be determined by rare, most conductive channels formed by random regions of relatively high carrier concentration, rather than the percolation cluster of mesh size L_c. Assuming that the resistors with $R = R_0 \exp(\xi)$ and $\xi < \xi_L$ are involved, the probability of finding the number $L/2r_c$ of such resistors forming a chain through the film between the electrodes can be written as $\exp[(L/2r_c)\ln(\xi_L/\xi_{max})]$. Dividing the latter by that chain resistance $R_0 \exp(\xi_L)$ gives the partial conductance of chains with $\xi \leq \xi_L$. Optimizing the exponent of the latter ratio with respect to ξ_L gives the optimum chain parameter $\xi_L = L/2r_c$. As a result, the conductance of the film can

be estimated as,

$$\sigma \propto \exp\left\{-\frac{L}{2r_c}\left[\ln\left(\frac{V_{max}}{kT}\frac{2r_c}{L}\right)+1\right]\right\}.$$

In the latter equation, one can impose the condition $\sigma = \sigma_\infty \equiv \sigma_0 \exp(-E_a/kT)$ when $L = L_c$, where σ_∞ has the meaning of the bulk conductivity. As a result, the effective conductivity of thin ($L < L_c$) structures can be written in the form,

$$\sigma = \sigma(F)\exp\left\{\frac{L_c - L}{2r_c}\left[\ln\left(\frac{V_{max}}{kT}\frac{2r_c}{L}\right)+1\right]\right\}, \tag{4.24}$$

where $\sigma(F)$ is given by Eq. (4.23). Here we have neglected the difference between logarithmic terms evaluated at L_c and L and have taken into account that $\xi_{max} = V_{max}/kT$, where V_{max} is the maximum transport barrier.

One prediction of Eq. (4.24) is that the effective activation energy of conduction $E_a = |d\ln(\sigma)/d(1/kT)|$ will decrease as the film thickness decreases below $L = L_c$. Another prediction refers to the case of extremely small devices with area below $A_c \sim r_c^2 \exp\{-(L/2r_c)\ln[(2r_c/L)(V_{max}/kT)]\}$ so that the above defined optimum channel is unlikely to be found within the device area. For such devices, conductance will be determined by the most efficient of the available channels, which will differ between samples; hence, there will be strong fluctuations in conductance between nominally identical devices. According to our rough estimates, that might occur well below the 10 nm scale.

4.3.8 Conduction Through Crystalline Inclusions in Amorphous Matrix

It is known that the reset pulse in chalcogenide PCM melts the material which then cools down fast enough to freeze in the amorphous phase, forming a dome (sometimes called a 'mushroom') as sketched in Fig 4-15. This melting-to-freezing transition is believed [18, 104] to result in a number of crystalline particles embedded

in the amorphous matrix. The latter scenario results in an interesting possibility that the system conductance will be governed by potential fluctuations created by the embedded crystallites. We note parenthetically that the presence of embedded crystallites follows from the standard thermodynamic consideration for the case of any glass possessing a crystalline counterpart of lower chemical potential.

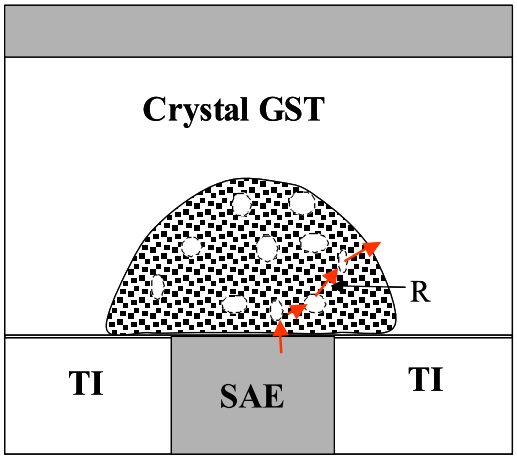

Figure 4-15: Amorphous dome with crystalline inclusions as part of the typical PCM structure including a small area electrode (SAE) and thermal insulator (TI). R is the average distance between crystallites. Arrows represent the current flow utilizing a path of minimum resistance.

Our model is based on the known valence band offset $\Delta \approx 0.4$ eV between the amorphous and crystalline phases (see Fig. 4-16). According to the standard principles of heterojunction physics, this offset is accommodated by the system through electrostatic screening. The screened potential is described by the standard Poisson equation $\nabla^2\phi = -4\pi\rho$ where the charge density is in turn related to the potential ϕ. That relation depends on the density of electron states $g(E)$, which, following the approach in Ref. [35], we assume constant. This gives $\rho = \phi q^2 g$ and the Poisson equation reduces to,

$$\nabla^2\phi = -\phi/r_s^2 \quad \text{with} \quad r_s = 1/\sqrt{4\pi q^2 g}, \tag{4.25}$$

where r_s has the physical meaning of the screening radius.

The solution of Eq. (4.25) for a spherically symmetric case is well known, $\phi \propto r^{-1}\exp(-r/r_s)$. The coefficient in front of it is determined by the boundary condition $q\phi(r_x) = \Delta$, where r_x is the crystallite radius. As a result each crystallite creates a potential,

$$\phi(r) = \Delta\frac{r_x}{qr}\exp\left(\frac{r_x - r}{r_s}\right) \quad \text{when} \quad r > r_x; \tag{4.26}$$

$r_s \gtrsim 100$ nm in the typical chalcogenide glasses. We note that the above assumption of constant density of states is not very restrictive as long as we are interested in distances shorter than r_s that is $\phi(r) \approx \Delta r_x/qr$; this can be readily verified for another standard case of a single-level density of states often used for crystalline semiconductors.

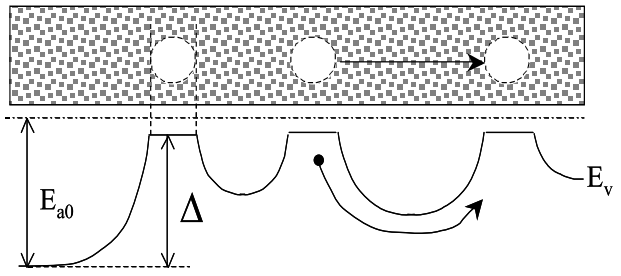

Figure 4-16: Top: a fragment of amorphous matrix with embedded crystallites. Bottom: energy band diagram showing valence band edge E_v in the crystalline and amorphous matrix (with offset Δ) and the activation energy E_{a0} is an amorphous phase without crystallites. Dot-dashed line represents the chemical potential. Arrows show the current flow between two crystallites.

The potential in Eq. (4.26) is the same as that of a coulombic center with effective charge,

$$Zq = \frac{r_x\varepsilon\Delta}{q} \sim 10q.$$

Therefore, one can use the entire wealth of results known for systems of charged centers in semiconductors to derive the following implications:

(1) Fluctuations of the electrostatic potential energy exist with the characteristic screening radius r_s and amplitude [85] $\delta U = Zq^2\sqrt{n_c r_s^3}/(\varepsilon r_s)$, where the square root represents the fluctuation in the number of charged crystallites of concentration n_c in a volume of radius r_s. Taking into account the above definition for Z, one can write,

$$\delta U \approx \Delta\sqrt{v\frac{r_s}{r_x}}, \tag{4.27}$$

where $v \equiv n_c r_x^3$ is the volume fraction occupied by crystalline particles. Using the above mentioned parameters it can be rather significant, $\delta U \gg kT$.

(2) The average decrease in the mobility edge, $\delta E_m \approx Zq^2(n_c)^{1/3}/\varepsilon$ representable as,

$$\delta E_m \approx \Delta v^{1/3}, \tag{4.28}$$

can be significant as well. The total decrease in activation energy of conductivity due to percolation can be estimated as,

$$\delta E_a = \delta E_m + \nu\delta U, \tag{4.29}$$

where ν is a numerical multiplier of order unity. It is dominated by its fluctuation component δU as long as the average distance between crystallites is shorter than the screening radius, $R < r_s$.

(3) The Poole-Frenkel effect appears here without any additional assumptions about the presence of coulombic centers in a material. The consideration in Sec. 4.3.1 will apply with corresponding renormalizations of the center's charge, $q \to Zq$. For example, the critical field of interplay between the regimes of one- and two-center field ionization regimes will become $Ze/4\varepsilon a^2$; numerically, it is $\sim 10^5$ V/cm when the

distance between crystallites is $a \approx 10$ nm. The two-center ionization effect results in the current,

$$I = I_0 \exp\left(\frac{\varepsilon r_x a F \Delta}{qkT}\right), \tag{4.30}$$

and the one-center effect becomes

$$I = I_0 \exp\left\{\frac{2\varepsilon}{kT}\sqrt{\left(\frac{r_x \Delta}{q}\right)^3 F}\right\}. \tag{4.31}$$

(4) All the implications of the percolation conduction mechanism in Sec. 4.3.7 will be applicable here. One specification is that the correlation length L_c [see Eq. (4.23)] for a system of charged particles becomes equal to the screening radius r_s. We note that for very small devices with size $L < r_s$, the size will play the role of screening radius.[105] In the latter case, Eq. (4.23) reduces to $\sigma(V) = \sigma(0) \exp\left(\eta\sqrt{qV/kT}\right)$.

Overall, the mechanism described in this subsection suggests the important role of the reset characteristics that determine the shape and composition of the amorphous dome in PCM devices. We shall briefly touch upon this issue further in Sec. 4.4 below.

4.4 Discussion

The non-ohmic conduction mechanisms described in this work are listed in Table 4.1 along with their characteristic relations and corresponding domains of applicability. Based on the experimental data, the only ones that can be excluded outright are Schottky emission and classical hopping conduction.

Shown in Fig. 4-17 are examples of typical data fitting corresponding to the various current voltage dependencies of the above discussed models; in each case one fitting equation was used to fit the entire IV curve. It appears that $\ln(I/I_0) \propto V$

provides the best fit, while the fit with $\ln(I/I_0) \propto -1/\sqrt{V}$ is the least successful. We note, however, that the approach implemented in Fig. 4-17 postulates a single dependence throughout the entire region of voltages. In reality, as we have seen, a single dependence fit may not be adequate: more than one non-ohmic domain with different temperature and thickness dependencies is typically observed, with a faster growing current in the pre-switching region. Eliminating the latter and limiting the single-curve fitting to relatively low voltages, the models (a)-(c) of Fig. 4-17 fit equally well, while the model (d) remains the least fitting, as illustrated in Fig. 4-18.

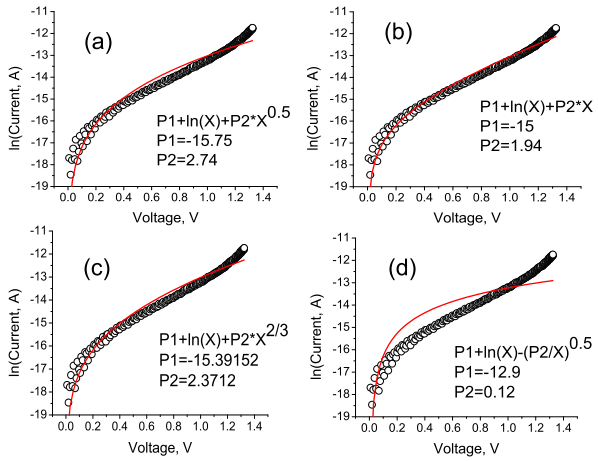

Figure 4-17: Four different fits of the same typical IV curve (presented also in Fig. 4-1) in the reset state of GST based PCM structure corresponding to the expressions discussed in the text: (a) $\ln(I/I_0) \propto \sqrt{V}$, (b) $\ln(I/I_0) \propto V$, (c) $\ln(I/I_0) \propto V^{2/3}$, and (d) $\ln(I/I_0) \propto -1/\sqrt{V}$.

These observations indicate that IV data fitting alone may not be conclusive enough to identify the most adequate model of transport in chalcogenide glasses. As can be seen in Table 4.1, several mechanisms provide the appropriate field dependence within an applicable domain. We note that several transport mechanisms may be appropriate with one mechanism prevailing in a particular domain of electric field, temperature, or thickness. For example, data for thickness dependence suggests that

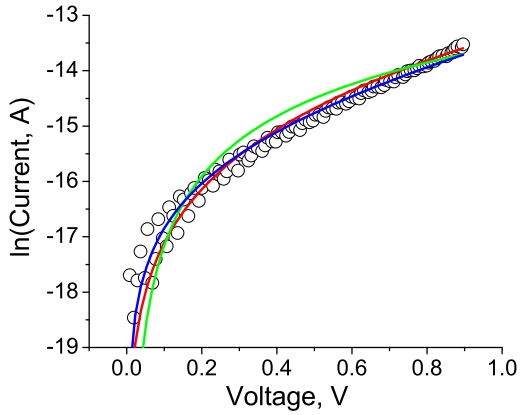

Figure 4-18: Fitting the data of Fig. 4-17 in the domain of $V < 1.0V$, which excludes the steep increase near threshold. The models with $\ln(I/I_0) \propto \sqrt{V}$ and $\ln(I/I_0) \propto V$, fit equally well, while that of $\ln(I/I_0) \propto -1/\sqrt{V}$ remains the worst fit.

space charge limited current may dominate below room temperature for samples of thickness $L > 1$ μm.[60, 64]

In our opinion, additional detailed studies are required to discriminate between the different mechanisms listed in Table 4.1. Such facts would include information about the temperature and size dependencies of conduction, resistance noise dependencies, statistical variations between different samples, differences between glasses with and without memory, and possibly some others.

An example of one such indicative fact is the observed dependencies [106] of the threshold voltage V_{th} and threshold current I_{th} on the *ohmic* resistance of the PCM reset state R_{res}, where V_{th} is logarithmic in R_{res} and I_{th} is the reciprocal of R_{res}. In some devices, those relations may be masked by geometrical or specific design related features. Yet, these results may point to the percolation nature of conduction discussed in Sec. 4.3.7 above and are applicable to spatially nonuniform systems, including the case of crystalline inclusions discussed in Sec. 4.3.8. Indeed, according to percolation theory the ohmic resistance is $R = R_0 \exp(E_a/kT)$, where E_a is the percolation transport barrier height. Assuming the barrier shape to be parabolic, its

linear dimension can be estimated as $l = r_c\sqrt{E_a/kT} = r_c\sqrt{\ln(R/R_0)}$, where r_c is the correlation radius (possibly equal to that of the medium range order). Because most of the voltage drops across that barrier, one can write $V_{th} = lF_{th} \propto \sqrt{\ln(R/R_0)}$, in qualitative agreement with the observations. On the other hand, the usual relation $I_{th} \propto \exp(-E_a/kT) \propto 1/R$ holds.

We shall end this section with a brief discussion of the case of extremely thin devices where optimum channel field emission described by Eq. (4.22) or percolation conduction in very thin films described by Eq. (4.24) can apply. Here, we use a possible example of such data from Ref. [67] aimed at studying thickness dependence in the limit of small amorphous volumes in GST PCM cells. From that data, we have plotted the resistance R as $\ln(R/R_0)$ vs. $\sqrt{L}$ for thicknesses between 8 and 35 nm. As show in Fig. 4-19, Eq. (4.22) provides good agreement with the data which is indicative of the optimum channel field emission mechanisms (see Sec. 4.3.6.2). However, as mentioned in Sec. 4.1, the thickness values were not directly measured but were inferred from IV measurements using a modified PF model. On the other hand, the amorphous thickness was varied by varying the reset pulse, which can also affect the crystal fraction in the resulting amorphous region. In that case, conduction through crystalline inclusions should be accounted for.

4.5 Conclusions

In conclusion, we have presented an overview of the experimental data and a comparative analysis of different d.c. transport mechanisms based on the established physics of chalcogenide glasses. Several models are capable of accounting for the various observed exponential field dependencies of the conductivity. Of the mechanisms considered, Poole-Frenkel ionization, field-induced delocalization of tail states, space-charge limited currents, optimum channel hopping in thin films, optimum channel field emission, percolation band conduction, and transport through crystalline inclu-

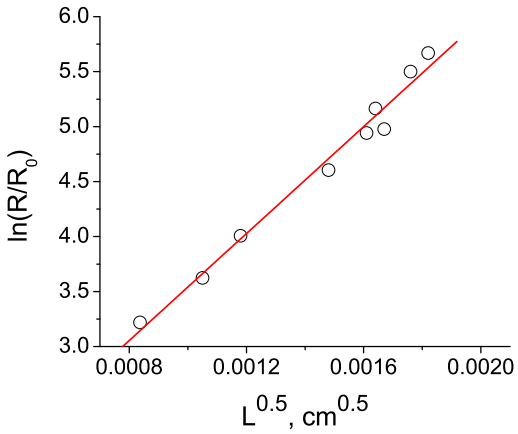

Figure 4-19: Logarithm of film resistance R vs. square root of thickness L. Data from Ref. [67] (circles) is fit using the optimum channel field emission mechanism of Eq. (4.22).

sions are all candidate explanations, while Schottky emission and classical hopping conduction are very unlikely.

We have shown that it is difficult to identify a particular mechanism through the analysis of IV data alone and further studies are required to discriminate between the different mechanisms.

Superimposed on the above listed mechanisms is a possibility of second phase conductive filament growing through the bulk under electric bias. If stable enough, such a filament will concentrate the electric field lines adding to the nonlinearilty of the measured conductivity.

We note the two original mechanisms of conductivity introduced in this Chapter: (1) more efficient than percolation the pinhole transversal conduction through the random potential relief of very thin films, and (2) conduction through the potential relief created by crystalline inclusions. The former is clearly in agreement with the data.

Because of the remaining uncertainty with the choice of the underlying mechanism

of bulk conduction, we can only suggest that more experimental study need to be conducted specifically aimed at establishing more indicative facts. More specifically we propose study of second phase nucleating vs. non-nucleating materials (i. e. PCM vs TS) materials - in order to check the hypothesis of crystalline inclusion facilitated conduction, as well as studying conduction under lower temperatures and correlating it with the material degree of disorder, and looking into size dependent effects and statistics of measurements between nominally identical samples.

Chapter 5

Summary and Conclusions

We present a model of data retention for phase change memory devices in which the active medium is a thin layer of chalcogenide glass. Data retention capability is compromised when a crystalline path is spontaneously formed in the glassy host, essentially shunting the device. We determine the probability and statistics of device failure for systems in which the crystalline volume fraction is below the critical volume fraction of percolation theory. In that regime, we show that rectilinear crystalline path formation is favored and we determine the criteria for when such paths dominate over the typical percolation cluster scenario. Our analytical approach, based on modeling the formation of such paths in terms of a half-space random walk, leads to closed form expressions that relate data retention characteristics to device parameters. The model is used to examine the effects of device geometry, temperature and external fields. The temporal statistics of device reliability are also considered for several failure mechanisms. A computer simulation is employed that supports our derived relationships between failure probability and device parameters.

We show that the average parameters of conductive filaments and the related characteristics of threshold switches can be described thermodynamically based on the system free energy. In particular, we derive analytical expressions for the filament radius as a function of applied bias, and its current-voltage characteristics, the observations of which have remained unexplained for about 30 years. Our description

is extendible to filament transients and allows for efficient numerical simulations of arbitrary switching structures.

Despite having been studied extensively studied over the last half century, the nature of the observed non-ohmic conduction in chalcogenide glasses is still under debate. An overview of the pertinent experimental data was followed by a review of the physics of localized states that are peculiar to chalcogenide glasses. We then described and evaluated twelve relevant transport mechanisms with conductivities that depend exponentially on the electric field. Most of the candidates provide more or less satisfactory fits of the observed non-linear IV data. Our analysis calls upon additional studies that would enable one to discriminate between the various alternative models.

References

[1] R. Bez, Tech. Dig. - Int. Electron Devices Meet. **2009**, 89.

[2] D. Kau, S. Tang, I. V. Karpov, R. Dodge, B. Klehn, J. A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee, T. Langtry, K. W. Chang, C. Papagianni, J. Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro, and G. Spadini, Tech. Dig. - Int. Electron Devices Meet. **2009**, 617.

[3] S. Lombardo, J. H. Stathis, B. P. Linder, K. L. Pey, F. Palumbo, C. H. Tung, J. Appl. Phys. **98**, 121301 (2005)

[4] J. H. Stathis, J. Appl. Phys., **86**, 5757 (1999).

[5] S.H. Demtsua, D.S. Albin, J.W. Pankow, A. Davies, Sol. Energy Mater. Sol. Cells **90**, 2934 (2006).

[6] T. J. McMahon, T. J. Berniard, and D. S. Albin, J. Appl. Phys. **97**, 054503 (2005).

[7] U. Russo, D. Ielmini, and A. L. Lacaita, IEEE Int. Reliab. Phys. Symp. Proc. **2007**, 547.

[8] U. Russo, D. Ielmini, A. Redaelli, and A. L. Lacaita, IEEE Trans. Electron Devices **53**, 3032 (2006).

[9] A. L. Efros and B. I. Shklovskii, *Electronic properties of doped semiconductors* (Verlag, Berlin, 1979).

[10] B. Gleixner, A. Pirovano, J. Sarkar1, F. Ottogalli, E. Tortorelli, M. Tosi, and R. Bez, IEEE Int. Reliab. Phys. Symp. Proc. **2007**, 542

[11] D. Mantegazza, D. Ielmini, A. Pirovano, A. L. Lacaita, IEEE Electron Device Lett. **28**, 865 (2007)

[12] M. Nardone, M. Simon, and V. G. Karpov, Appl. Phys. Lett., **96**, 163501 (2010).

[13] M. Pollak and J. J. Hauser, Phys. Rev. Lett. **31**, 21 (1973).

[14] M. E. Raikh and I. M. Ruzin, in *Mesoscopic Phenomena in Solids*, edited by B. L. Altshuller, P. A. Lee, and R. A. Webb (Elsevier, 1991), p. 315.

[15] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers 2d Edition* (Dover Publications, 2000)

[16] M. Slutsky, Am. J. Phys. **73**, 308 (2005).

[17] I. V. Karpov and S. A. Kostylev, IEEE Electron Device Lett., **27**, 808 (2006)

[18] Y.H. Shih, M.H. Lee, M. Breitwisch, R. Cheek, J.Y. Wu, B. Rajendran, Y. Zhu, E.K. Lai, C.F. Chen, H.Y. Cheng, A. Schrott, E. Joseph, R. Dasaka, S. Raoux, H.L. Lung, and C. Lam, Tech. Dig. - Int. Electron Devices Meet. **2009**, 753

[19] J. A. Kalb, C. Y. Wen, and Frans Spaepen, J. Appl. Phys. **98**, 054902 (2005).

[20] L.D. Landau, I. M. Lifshits, *Electrodynamics of continuous media* (Pergamon, Oxford, New York, 1984).

[21] V. G. Karpov, Y. A. Kryukov, I. V. Karpov, and M. Mitra, J. Appl. Phys. **104**, 054507 (2008). I. V. Karpov, M. Mitra, G. Spadini, U. Kau, Y. A. Kryukov, and V. G. Karpov, J. Appl. Phys. **102**, 124503 (2007).

[22] F. Bedeschi, R. Fackenthal, C. Resta, E. M. Donze, M. Jagasivamani, E. C. Buda, F. Pellizzer, D. W. Chow, A. Cabrini, G. Calvi, R. Faravelli, A. Fantini,

G. Torelli, D. Mills, R. Gastaldi, and G. Casagrande, IEEE Journal of Solid-State Circuits **44**, 217 (2009).

[23] D. Adler, H.K. Henisch, and S.N. Mott, Rev. Mod. Phys. **50**, 209 (1978).

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

[25] K. E. Petersen and D. Adler, J. Appl. Phys. **47**, 256 (1976).

[26] B. K. Ridely, Proc. Phys. Soc. **82**, 954 (1963).

[27] J. Ross, *Thermodynamics and Fluctuations Far from Equilibrium* (Springer, New York, 2008), p. 119.

[28] A. Redaelli, A. Pirovano, A. Benvenuti, and A. L. Lacaita, J. Appl. Phys. **103**, 11 (2008).

[29] K. E. Petersen and D. Adler, J. Appl. Phys. **50**, 5065 (1979).

[30] I. V. Karpov, M. Mitra, D. Kau, G. Spadini, A. Y. Kryukov, and V. G. Karpov, Appl. Phys. Lett. **92**, 173501 (2008); V. G. Karpov, Y. A. Kryukov, I. V. Karpov, and M. Mitra, Phys. Rev. B **78**, 052201 (2008).

[31] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Elsevier, Amsterdam, Boston, 2008).

[32] M. C. Weinberg and G. F. Nelson, J. Non-Cryst. Solids **74**, 177 (1985); C. Barrett, W. Nix, and A. Tetelmam, *The Principles of Engineering Materials* (Prentice-Hall, Englewood Cliffs, NJ, 1973); X. S. Miao, L. P. Shi, H. K. Lee, J. M. Li, R. Zhao, P. K. Tan, K. G. Lim, H. X. Yang, and T. C. Chong, Jpn. J. Appl. Phys., Part 1 **45**, 3955 (2006).

[33] A. E. Owen and J. M. Robertson, IEEE Trans. Electron Devices **ED-20**, 105 (1973).

[34] D. L. Kencke, I. V. Karpov, B. G. Johnson, Sean Jong Lee, DerChang Kau, S. J. Hudgens, J. P. Reifenberg, S. D. Savransky, Jingyan Zhang, M. D. Giles, G. Spadini Tech. Dig. - Int. Electron Devices Meet. **2007**, 323.

[35] N. F. Mott and E. A. Davis, *Electronic Processes in Non-crystalline Materials* (Clarendon Press, Oxford, 1979).

[36] S. Lai and T. Lowrey, Tech. Dig. - Int. Electron Devices Meet. 2001, 803.

[37] S. Eilert, M. Leinwander, G. Crisenza, in *Proceedings of the IEEE International Memory Workshop, IMW 2009*, Monterey, CA, 2009 (IEEE, New York, 2009), p.1-2.

[38] T. Morikawa, K. Kurotsuchi, M. Kinoshita, N. Matsuzaki, Y. Matsui, Y. Fujisaki, S. Hanzawa, A. Kotabe, M. Terao, H. Moriya, T. Iwasaki, M. Matsuoka, F. Nitta, M. Moniwa, T. Koga, and N. Takaura, Tech. Dig. - Int. Electron Devices Meet. 2007, 307.

[39] J. Rodgers, J. Maimon, T. Storey, D. Lee, M. Graziano, L. Rockett, K. Hunt, in *Proceedings of the IEEE Nonvolatile Memory Technology Symposium, 2008 NVMTS*, Pacific Grove, CA, 2008 (IEEE, New York, 2008), p. 1.

[40] A. Fazio, Tech. Dig. - Int. Electron Devices Meet. 2009, 641; G. Burr, M. Breitwisch, M. Franceschini, D. Garetto, K. Gopalakrishnan, B. Jackson, B. Kurdi, C. Lam, L. A. Lastras, A. Padilla, B. Rajendran, S. Raoux, R. S. Shenoy, J. Vac. Sci. Technol. **B28**, 223 (2010).

[41] F. Bedeschi, R. Bez, C. Boffino, E. Bonizzoni, E. C. Buda, G. Casagrande, L. Costa, M. Ferraro, R. Gastaldi, O. Khouri, F. Ottogalli, F. Pellizzer, A. Pirovano, C. Resta, G. Torelli, M. Tosi, IEEE J. Solid-St. Circ. **40**, 1557 (2005).

[42] J. H. Oh, J. H. Park, Y. S. Lim, H. S. Lim, Y. T. Oh, J. S. Kim, J. M. Shin, J. H. Park, Y. J. Song, K. C. Ryoo, D. W. Lim, S. S. Park, J. I. Kim, J. H. Kim,

J. Yu, F. Yeung, C. W. Jeong, J. H. Kong, D. H. Kang, G. H. Koh, G. T. Jeong, H. S. Jeong, and K. Kim, Tech. Dig. - Int. Electron Devices Meet. 2006, S2P6.

[43] F. Pellizzer, A. Benvenuti, R. Gleixner, Y. Kim, B. Johnson, M. Magistretti, T. Marangon, A. Pirovano, R. Bez, and G. Atwood, in *Proceedings of the Symposium on VLSI Technology, 2006*, Honolulu, HI, 2006 (IEEE, New York, 2006), p. 122.

[44] G. Servalli, Tech. Dig. - Int. Electron Devices Meet. 2009, 113.

[45] D. C. Kau, S. Tang, I. V. Karpov, R. Dodge, B. Klehn, J. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee, T. Langtry, K. 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.

[46] I. V. Karpov, D. Kencke, D. Kau, S. Tang, and G. Spadini, in *Proceedings of the Materials Research Society Symposium* San Francisco, PA, 2010 (MRS, Warrendale, PA, 2010), Vol. 1250, G14-01-H07-01.

[47] R. M. Hill, Phil. Mag. **23**, 59 (1971).

[48] J. Frenkel, Phys. Rev. **54**, 657 (1938).

[49] H. H. Poole, Lond. Edinb. Dubl. Phil. Mag. **33**, 112 (1916); *ibid.,* **34**, 195 (1917).

[50] T. Gotoh, J. Non-Cryst. Solids **353**, 2728 (2008).

[51] V. K. Saraswat, V. Kishore, Deepika, N. S. Saxena, T. P. Sharma, L. I. Singh, and P. K. Saraswat, Chalcogenide Letters **5**, 95 (2008).

[52] M. M. El-Samanoudy, Appl. Surface Science **207**, 219 (2003).

[53] E. A. Lebedev, S. A. Kozykhin, N. N. Konstantinova, and L. P. Kazakova, Semiconductors **43**, 1343 (2009).

[54] N. A. Bogoslowsky and K. D. Tsendin, Semiconductors **43**, 1338 (2009).

[55] E. N. Voronkov and S. A. Kozyukhin, Semiconductors **43**, 921 (2009).

[56] D. Ielmeni and Y. Zhang, J. Appl. Phys. **102**, 054517 (2007); Appl. Phys. Lett. **90**, 192102 (2007); D. Ielmini, Phys. Rev. B **78**, 035308 (2008).

[57] D. Allsopp and M. J. Thompson, J. Phys. D: Appl. Phys. **9**, 2075 (1976).

[58] E. A. Lebedev and N. A. Rogachev, Sov. Phys. Semiconductors **15**, 86 (1981).

[59] M. P. Shaw, S. H. Holmberg. and S. A. Kostylev, Phys. Rev.Lett. **31**, 542 (1973).

[60] I. F. Kodgespirova, V. A. Shkut, and S. A. Kostylev, in *Proceedings of the International Conference on Amorphous Semiconductors, Bucharest, 1982*, in Russian, edited by M. Popesku, p. 247-249.

[61] S. D. Savransky and I. V. Karpov, in *Phase-Change Materials for Reconfigurable Electronics and Memory Applications*, edited by A. H. Edwards, P. J. Fons, S. Raoux, P. C. Taylor, and M. Wuttig (Mater. Res. Soc. Symp. Proc. **Volume 1072E**, Warrendale, PA, 2008), 1072-G06-09.

[62] J. J. Hauser and R. S. Hutton, Phys. Rev. Lett. **37**, 868 (1976).

[63] P. J. Walsh, R. Vogel, and E. J. Evans, Phys. Rev. **178**, 1274 (1969).

[64] J. M. Robertson, Ph.D. thesis, University of Edinburgh, 1971.

[65] H. J. de Wit and C. Crevecoeur, J. Non-Cryst. Solids **10**, 787 (1972).

[66] F. Pellizzer, A. Pirovano, F. Ottogalli, M. Magistretti, M. Scaravaggi, P. Zuliani, M. Tosi, A. Benvenuti, P. Besana, S. Cadeo, T. Marangon, R. Moranti, R. Piva, A. Spandre, R. Zonca, A. Modelli, E. Varesi, T. Lowrey, A. Lacaita, G. Casagrande, and R. Bez, in *Proceedings of the Symposium on VLSI Technology, 2004* (IEEE, New York, 2004), p. 1819.

[67] D. Fugazza, D. Ielmini, S. Lavizzari, and A. L. Lacaita, Tech. Dig. - Int. Electron Devices Meet. 2009, 1.

[68] A. Calderoni, M. Ferro, D. Ielmini, and P. Fantini, IEEE Electron Device Lett. **31**, 1023 (2010).

[69] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[70] P. W. Anderson, Phys. Rev. Lett. **34**, 952 (1975).

[71] S. D. Baranovskii and V. G. Karpov, Fiz. Tech. Poluprov. **21**, 3 (1987) [Sov. Phys. Semicond. **21**, 1 (1987)].

[72] S. D. Baranovskii and V. G. Karpov, Fiz. Tech. Poluprov. **21**, 314 (1987) [Sov. Phys. Semicond. **21**, 189 (1987)].

[73] R. A. Street and N. F. Mott. Phys. Rev. Lett. **35**, 1293 (1975).

[74] M. Kastner, D. Adler, and H. Fritzsche, Phys. Rev. Lett. **37**, 1504 (1976); M. Kastner and H. Fritzsche, Phil. Mag. **37**, 199 (1978).

[75] M. I. Klinger and V. G. Karpov, Zh. Eksper. Teor. Fiz. **82**, 1687 (1982) [Sov. Phys. JETP **55**,976 (1982)]; V. G. Karpov, *ibid.* **85**, 1017 (1983) [Sov. Phys. JETP **58**, 592 (1983)].

[76] V. G. Karpov, M. I. Klinger, and F. N. Ignatiev, Zh. Eksp. Teor. Fiz. **84**, 761 (1983) [Sov. Phys. JETP **57**, 439 (1983)].

[77] Yu. M. Galperin, V. G. Karpov, and V. I. Kozub, Adv. Phys. **38**, 669 (1989).

[78] D. A. Parshin, Sov. Phys. Solid. State **36**, 991 (1994); D. A. Parshin, Phys. Scr. **T49A**, 180 (1993); D. A. Parshin, H. R. Schober, and V. L. Gurevich, Phys. Rev. B **76**, 064206 (2007); B. Ruffle, D. A. Parshin, E. Courtens, and R. Vacher, Phys. Rev. Lett. **100**, 015501 (2008).

[79] W. A. Phillips, Phil. Mag. **34**, 983 (1976).

[80] V. G. Karpov, Fiz. Tekh. Poluprovdn. **19**, 123 (1984) [Sov. Phys. Semicond. **19**, 74 (1984)].

[81] A. R. Long. Adv. Phys. **31**, 553 (1982).

[82] M. Nardone, V. I. Kozub, I. V. Karpov, and V. G. Karpov, Phys. Rev. B **79**, 165206 (2009).

[83] V. Abakumov, V. Perel, and I. Yassievich, *Nonradiative Recombination in Semiconductors*, Modern Problems in Condensed Matter Science (North-Holland, Amsterdam, 1991), Vol. 33.

[84] S. Lee, D. S. Jeong, J. Jeong, W. Zhe, Y-W Park, H-W Ahn, B. Cheong, Appl. Phys. Lett. **96**, 023501 (2010).

[85] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, New York, 1992).

[86] V. I. Koldyaev, Phil. Mag. B **79**, 331 (1999).

[87] B.-S. Lee, J. R. Abelsona, S. G. Bishop, D-H. Kang and B. Cheong, K-B. Kim, J. Appl. Phys. **97**, 093509 (2005); M. Frumar, T. Wagner, M. Hrdlicka, B. Frumarovaa, and P. Nemec, E/PCOS (2005).

[88] S. M. Sze, *Physics of Semiconductor Devices* (Willey & Sons, New York, 1981).

[89] I. M. Lifshits, S. A. Gredeskul, and L. A. Pastur. *Introduction to the Theory of Disordered Systems* (John Wiley & Sons, New York, 1988).

[90] S. R. Elliott, J. Non-Cryst. Solids **97**, 159 (1987).

[91] M. A. Lampert and P. Mark, *Current Injection in Solids* (Academic Press, New York, 1970).

[92] I. V. Karpov, (unpublished).

[93] T. G. M. Kleinpenning, Physica B+C **94**, 141 (1978).

[94] J. M. Marshall and A. E. Owen, Phil. Mag. **31**, 1341 (1975).

[95] W. A. Phillips, Phil. Mag. **34**, 983 (1976).

[96] M. L. Knotek, Solid State Commun. **17**, 1431 (1975).

[97] B. Bagley, Solid State Commun. **8**, 345 (1970).

[98] S. D. Baranovskii and V. G. Karpov, Fiz. Tekh. Poluprovodn. **20**, 1811 (1986) [Sov. Phys. Semicond. **20**, 1137 (1986)].

[99] M. E. Raikh and I. M. Ruzin, in *Mesoscopic Phenomena in Solids*, edited by B. L. Altshuller, P. A. Lee, and R. A. Webb (Elsevier, New York, 1991), p. 315.

[100] E. I. Levin, I. M. Ruzin, and B. I. Shklovskii, Sov. Phys. Semicond. **22**, 401 (1998).

[101] M. Pollak and J. J. Hauser, Phys. Rev. Lett. **31**, 21 (1973).

[102] H. Fritzsche, J. Phys. and Chem. of Solids **68**, 878 (2007).

[103] B. I. Shklovskii, Sov. Phys. Semicond. **13**, 53 (1979).

[104] J. González-Hernández, E. F. Prokhorov, Yu. V. Vorobiev, E. Morales-Sánchez, A. Mendoza-Galván, S. A. Kostylev, Yu. I. Gorobets, V. N. Zakharchecnko, and R. V. Zakarchenko, J. Vac. Sci. Technol. A **19**, 1623 (2001); E. Morales-Sánchez, J. González-Hernández, and E. F. Prokhorov, J. Optoelectron. Adv. M. **3**, 333 (2001).

[105] M. Nardone, V. G. Karpov, D. C. S. Jackson, and I. V. Karpov, Appl. Phys. Lett. **94**, 103509 (2009).

[106] S. A. Kostylev, in *Proceedings of the Non-Volatile Memory Symposium* Portland, OR, 2009 (IEEE, New York, 2009), p. 16-23.

Appendix A

Derivations Related to the Shunting Probability

The probability distribution function for a 3-D unconstrained random walk to begin at the point $\mathbf{r}_0$ and end at the point $\mathbf{r}$ is given by

$$G = \left(\frac{3}{2\pi d^2 N}\right)^{3/2} e^{-3|\mathbf{r}-\mathbf{r_0}|^2/2d^2N} \tag{A.1}$$

where d is the average displacement for each step in the walk and N is the number of steps. If the walk is restricted to the upper half-space ($z > 0$) then we may use the method of images to find the resulting probability distribution. Assume the random walk starts at some small distance a above the $x - y$ plane ($\mathbf{r}_0 = a\hat{\mathbf{z}}$). Then

$$G = \left(\frac{3}{2\pi d^2 N}\right)^{3/2} \left(e^{-3|\mathbf{r}-a\hat{\mathbf{z}}|^2/2d^2N} - e^{-3|\mathbf{r}+a\hat{\mathbf{z}}|^2/2d^2N}\right) \tag{A.2}$$

$$= 2\left(\frac{3}{2\pi d^2 N}\right)^{3/2} e^{-3(r^2+a^2)/2d^2N} \sinh\frac{3az}{d^2N} \tag{A.3}$$

which gives the required $G(z = 0) = 0$. Assuming the walk ends sufficiently far from the starting point ($r^2 \gg a^2$) and the number of steps is large ($d^2N \gg az$) then the above may be approximated as

$$G = \frac{az}{\sqrt{2\pi^3}} \left(\frac{3}{d^2 N} \right)^{5/2} e^{-3r^2/2d^2N} \tag{A.4}$$

When the method of images was used, the distribution function was no longer normalized. Since we require the walk to end somewhere in the upper half-space, the normalized probability distribution G' is then

$$G' = \frac{G}{\int_{z>0} G \; dxdydz} = \frac{z}{2\pi} \left(\frac{3}{d^2 N} \right)^2 e^{-3r^2/2d^2N} \tag{A.5}$$

The result above is the general form for the probability distribution for an unconstrained random walk in half space. The probability that the walk will end at the opposite electrode will depend on the particular geometry involved (see Fig. 2-6).

The simplest case is when the other electrode is a plane parallel to the starting plane a distance L away [Fig. Fig. 2-6 (a)]. Integrating Eq. (A.5) over the ending plane gives the corresponding probability distribution p_z. By definition, the probability of ending in a slab of thickness δz a distance L away is given by $p_z(L)\delta z$. By setting the thickness of the slab equal to the lattice spacing d, we have

$$p_N = p_z(L)d = \int_{\infty}^{\infty} \int_{\infty}^{\infty} \frac{L}{2\pi} \left(\frac{3}{d^2 N} \right)^2 e^{-3(x^2+y^2+L^2)/2d^2N} \; dxdy \tag{A.6}$$

$$= \frac{3L}{Nd} e^{-3L^2/2d^2N} \tag{A.7}$$

which is the same as Eq. (2.4) when the lattice spacing is set to the nucleation diameter ($d = 2R$). Eqs. (2.11) and (2.11) are obtained in a similar manner.

Appendix B

Approximations Used in the Analytical Modeling of Stable Filament Radius

Assuming the filament is sufficiently wide ($r \gtrsim h$), the current density will be nearly uniform and concentrated within the filament. As such, the temperature increase due to Joule heating occurs mostly within the filament and the integral for the thermal contribution may be approximated by the average temperature increase within the filament ΔT multiplied by the filament volume. Similarly, since the field within a parallel plate capacitor is approximately uniform and perpendicular to the surface, the integral for the electrostatic contribution is proportional to the magnitude of the field squared. From Eq. (3.3) the free energy is then given by,

$$F = C_v \Delta T \pi r^2 h + \frac{E^2 \varepsilon}{8\pi} A h + 2\pi r h \sigma + \pi r^2 h \mu. \tag{B.1}$$

Treating the external load resistance and device as resistors in series (see Fig. 3-1), the voltage across the device is given by $V_d = V/(1 + R_L/R)$ where V is the source voltage, R_L is the load resistance, and R is the resistance of the device. The

magnitude of the electric field a given by the voltage divided by thickness, and so

$$E = \frac{V}{h}\left(\frac{R}{R+R_L}\right) \tag{B.2}$$

The filament is connected in parallel to the surrounding amorphous material. Since the filament resistivity is much lower than in the amorphous phase, the device resistance is approximately the same as the filament resistance when $r^2 \gg A\rho_c/\rho_a \approx 10^{-4}$ A. If this is the case, the device resistances is given approximately by

$$R \approx \frac{\rho h}{\pi r^2} \tag{B.3}$$

The thermal conductivity of the electrodes is much greater than that of the amorphous phase. As such, the heat flow will be directed towards the electrodes. The temperature distribution within the filament (assuming a uniform current density as discussed previously) is found by solving the heat conduction equation $\chi\nabla^2 T = -\rho J^2$ where ρ is the filament electrical resistivity, and χ is the thermal conductivity taken to be the same for the filament and host materials. The solution is given by

$$T(z) = \frac{I^2\rho}{2\pi^2\chi r^4}(z - h/2)^2 \tag{B.4}$$

where I is the current. Integrating z from $-h/2$ to $h/2$ gives the average temperature increase

$$\Delta T_{avg} = \frac{I^2 h^2 \rho}{12\pi^2\chi r^4} \tag{B.5}$$

We noticed from our numerical simulations (see Sec. 3.2.4) that the thermal contribution to the free energy is better approximated by using the maximum temperature increase, which occurs at the center of the filament ($z = 0$). From Eq. (B.4), the maximum temperature increase is given by $3\Delta T_{avg}/2$.