Politecnico di Milano
Dipartimento di Elettronica, Informazione e Bioingegneria
Doctoral Program in Information Technology

Characterization and modeling of current transport and polarization switching in FTJs

Doctoral Dissertation of:
Giulio Franchini

Supervisor:
Prof. Alessandro Sottocornola Spinelli

Tutor:
Prof. Ivan Rech

The Chair of the Doctoral Program:
Prof. Barbara Pernici

2021 – XXXIV

Abstract

In the last decades PCs, smartphones and wearable and interconnected electronic gadgets are gaining more and more momentum. Due to this large success, the semiconductor market experienced an enormous growth, especially in the memory sector, as new devices require large amount of solid-state memory with high performance, both in terms of integration density and throughput. Since they are both needed in the aforementioned devices, such noticeable growth has affected both volatile and non volatile memories. In order to keep up with the non-volatile memories market requests, Flash memory technologies have been the object of an uninterrupted scaling process which led to increase their storage density, and let them become the most successful solution in the non-volatile memory landscape; the smallest feasible feature size, equal to 14 nm, was conceived around the middle of the 2010's decade. After hitting this limit, the scaling approach switched to an equivalent one, consisting in stacking memory one on top of the other exploiting the third dimension. Although this switch determined a general improvement in terms of reliability, the novel architecture of the memories brought along some new issues. In the volatile memory landscape the DRAM technology as well underwent an intensive scaling process to reduce the cost for stored bit, and en-

countered similar reliability issues. Such issues do not constitute the only obstacle to memory technologies. Another major challenge has to be dealt with: the so-called von Neumann bottleneck. In the last forty years, the performance gap between the central processing unit (CPU) and the working memory (DRAM) has never stopped growing. This gap results in a memory bottleneck that reduce the overall performance of a computing system.

The aforementioned issues fueled the research on novel memory technologies, both volatile and non-volatile, different from the ones that dominate the market. Among them we find: resistive RAM (RRAM), phase change memory (PCM), magnetoresistive random access memory (MRAM), and a wide variety of ferroelectric memories. This thesis focuses on Ferroelectric Tunnel Junction (FTJ), one of the most promising candidates of the novel foerroeletric memory technologies. In particular, an experimental characterization on FTJ samples has been carried out exploring the main features of the device: resistive window, retention, I-V characteristics. Experimental activities are followed by the development of a one-dimensional in-house MATLAB code for the device simulation, capable of reproducing the I-V characteristics and explaining some of the physical phenomena involved. During the last year of the Ph.D. efforts have been directed toward the analysis of the switching process, in an attempt to develop a two-dimensional switching simulator to be integrated with the 1-D one. The first two chapters of this thesis are introductory, and are intended to give all the background information needed to go through the remaining part of the Ph.D. thesis.

Chapter 1 provides a general introduction to memory technologies. The Flash and DRAM technologies are presented with a description of the issues related to their scaling process and some pursued solutions. At the end of the chapter the von Neumann bottleneck is briefly explained and some of the new memory technologies that could tackle the problem

are presented.

In Chapter 2 an introduction on the concept of polarization is given, followed by some non-exhaustive information about ferroelectricity and ferroelectric materials. Later in the chapter the most promising ferroelectric memories are presented: FeRAM, FeFET and FTJ. The last one is then illustrated in detail together with the meaningful equations and its performances.

Since it focuses on the experimental characterization and one-dimensional modeling, chapter 3 is the most dense section of this thesis. The experimental setup and the samples are presented, followed by all the measures and the setup changes needed. The following subsection is devoted to a full detailed explanation of the developed 1-D model, presenting the equations and following their implementation. In the last section of the chapter, measurements and simulations are compared, and physical phenomena analysed.

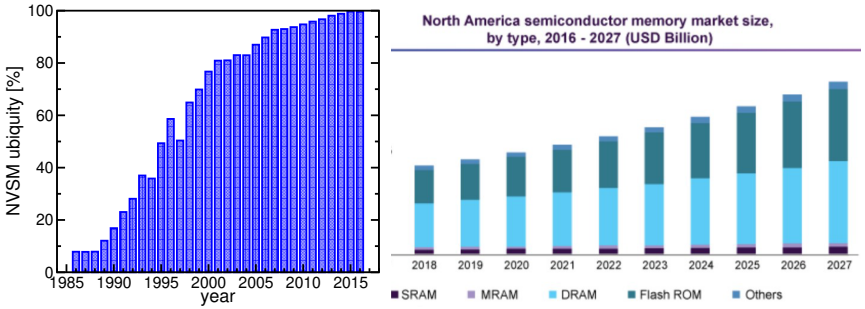
In Chapter 4 an introduction on ferroelectric switching is given, followed by some of the numerous analytical approaches employed to study it and by the presentation of the Landau-Devonshire-Ginzburg theory employed in the developed two-dimensional simulator. Such simulator is presented both from the analytical point of view and its MATLAB implementation. Finally, early results are presented together with the issues encountered and assumptions on how to overcome them.

Contents

Chapter 1

Memory devices

This chapter provides a general introduction to memory devices, dividing them in two categories, that is to say non-volatile and volatile memories, depending on their ability of retaining written data. The former operates at higher biases and is slower at writing and reading data but can retain them up to ten years even in the absence of power supply, where the latter is much quicker in writing and reading the data but cannot keep them if lacking power supply. An overview on the memory market will be given, briefly discussing their limitations and reasons to research new emerging technologies. Among them we find ferroelectric memories, which have witnessed a renewed interest in recent years due to the discovery of ferroelectric hafnium-oxide, a CMOS compatible material. The main ferroelectric memories structures are: FeRAM, FeFET and FTJ. They will be presented with their advantages and disadvantages related to the specific application they target. FTJs will be only briefly discussed here, as their working principles will be the main topic of the next chapter.



(a) **(b)**

Figure 1.1: (a) *Non-Volatile Semiconductor Memories* (NVSM) ubiquity in electronic systems (adapted from [1]) and (b) increase in memory market size for both volatile and non-volatile memories [2].

1.1 Memories growth and classification

The semiconductor market experienced an enormous growth in the last decades due to the extensive mass success of modern electronic equipment. PCs, smartphones and wearable and interconnected electronic gadgets are gaining more and more momentum in the last years. The fact that these devices require large amount of solid-state memories with high performance both in terms of integration density and throughput, made the memory market share of the semiconductor industry fairly relevant, with a revenue of about 128 billions, for more than 30% of its total revenues in 2017 [3]. Memory devices can be divided into two main categories depending on their ability to retain data for a long period of time when the power supply is switched off. These two categories are:

- **Non-Volatile Memories**: if, independently of the state of the power supply, information is stored for extremely long times ($\approx$ 10 years) .
- **Volatile Memories**: if, once the power is switched off, information is completely lost (although it can be stored

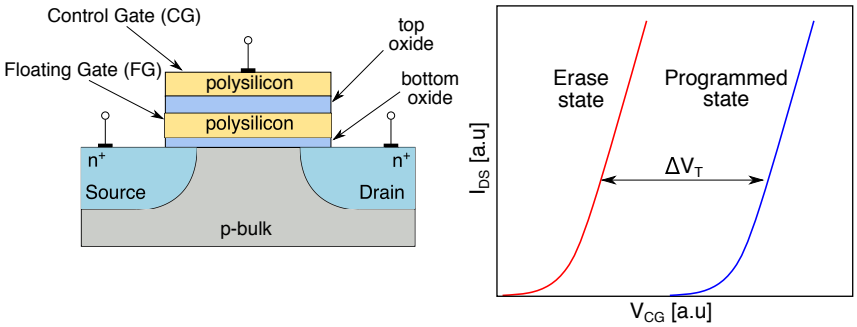


Figure 1.2: Schematic representation of a memory transistor adopting (a) a polysilicon floating gate (FG) and (b) the physical principle at the basis of the storing of information in a memory transistor: the V_T-shift experienced by the memory transistor is proportional to the amount of stored charge (with negative sign), determining two different states of the device (bit).

for short or long times when the system is on, like in the case of DRAM and SRAM memories, respectively).

The increasing diffusion of portable equipment for work and leisure time came together with the need of low cost and low power Non-Volatile Semiconductor Memories (NVSM), that enable the power down of such devices during idle periods without any loss of information. The number of electronic systems requiring some kind of NVSM skyrocketed in the last decades as shown in Fig. 1.1(a). This fact fueled the research to develop high-performance and reliable memories.

On the other hand, the volatile memories also experienced a dramatic growth, as pointed out in 1.1(b), and the research has been propelled in the same way toward higher performaces and new technologies and solutions.

1.2 Non-volatile memories

Non-volatile memories are a fundamental part of almost any electronic system. Compared to the volatile ones, non-volatile require higher powers and biases and have longer operation times. In the last decade, the Flash NAND technology emerged as the best solution for non-volatile solid-state memory. From its market introduction in 1987, this technology has tremendously grown to the point that nowadays the market share is about 49 billions USD worldwide and its annual growth rate is 52%. Due to the increasing demands from Big Data, IA and Iot, further growth is expected in the next years.

1.2.1 Memory transistor

The fundamental element at the basis of Flash memories is the so called *memory transistor*, which was invented in 1984 [4]. The main idea is to change the threshold-voltage (V_T) of a MOS structure, which is easily detected by varying the fixed charge stored in its gate insulator. The V_T depends not only on the magnitude of the stored charge but also on its spatial distribution. To better control this distribution in between the gate (usually referred to as *Control Gate*, CG, in this context) and the channel of a MOS transistor (usually n-type), a further layer is introduced so that the stored charge is confined at a fixed distance between substrate and gate. The most widespread solution is to employ a highly-doped polysilicon layer (see Fig. 1.2(a)) called *floating gate* (FG). In Fig. 1.2(a), we can see two different oxide layers, the top one being called blocking oxide and the lower one tunneling oxide. From simple electrostatics considerations [5], it can be shown that:

$$\Delta V_T = V_T - V_{T,0} = -\frac{Q_s}{C_{pp}} \tag{1.1}$$

where $V_{T,0}$ is the neutral V_T, i.e. when no net charge is present in the storage layer, and C_{pp} is the CG-to-FG capac-

itance. We can clearly see from Eq. 1.1 that the sign of ΔV_T is opposite to the sign of the net charge Q_s such that when Q_s is positive ΔV_T is negative and viceversa. Consequently, a proper discretization of Q_s can be binary-encoded, making the memory transistor store 1 bit of information. To *read* the stored information the exponential relation between ($V_{CG} - V_T$) and the drain-to-source current (I_{DS}) is exploited, as can be seen from Fig. 1.2(b). To discriminate between the two states is sufficient to measure I_{DS} at V_{CG} intermediate between the two V_T levels. Conventionally, the high-V_T cell state is indicated as *programmed*, whereas the low-V_T one is indicated as *erased*.

1.2.2 Cell operations

Any non-volatile memory storage must give the possibility to modify the stored information. In the case of the aforementioned Flash memory technology this means being able to change the stored charge Q_s. Many different physical mechanisms have been proposed and tested in order to modify Q_s. By far the most exploited are: *Fowler-Nordheim* (FN) tunneling and *Channel Hot Electron Injection* (CHEI). Both mechanisms produce an electron current across the bottom oxide, the difference between them is in the travelling electrons energy. In the case of FN tunneling, the charge exchange between storage layer and substrate is accomplished by quantum-mechanical tunneling of low-energy electrons through the potential barrier between the silicon channel and the bottom-oxide [6]. In the case of CHEI, electrons must have much higher kinetic energies because they have to surpass the same potential barrier. To do this, they are accelerated in the direction of the channel (hence they become *hot*) and have a finite probability to be swept toward the storage layer as a consequence of a scattering event with the lattice (this process, of course, benefits from vertical electric field in the direction of the CG, see [7] for a quantitative discussion).

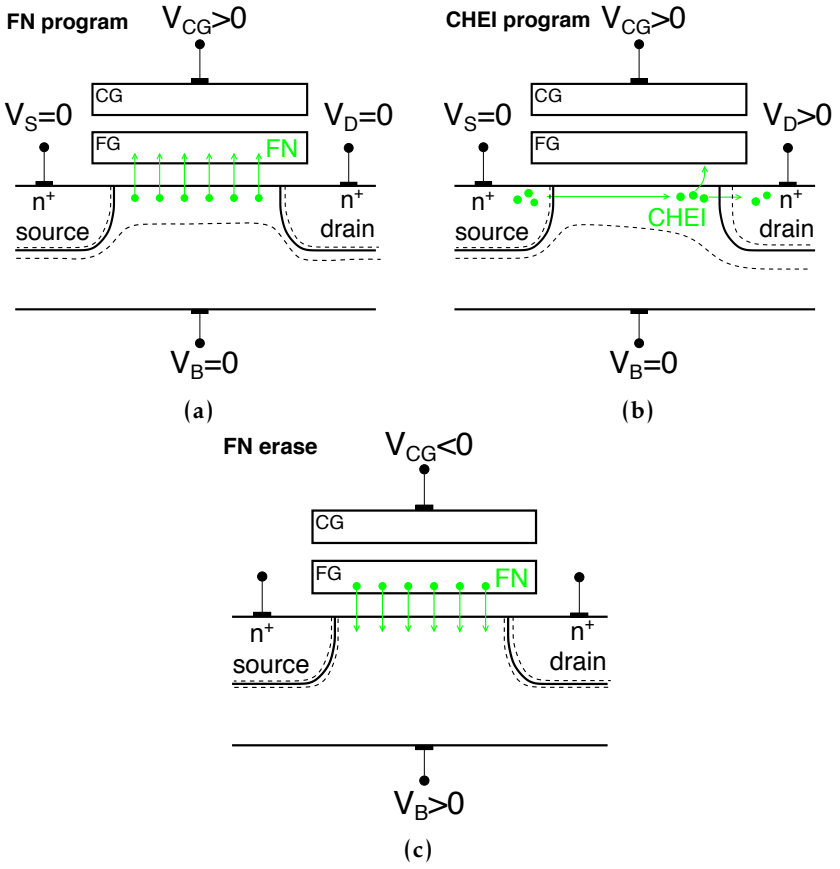


Figure 1.3: Voltages schemes adopted in a FG Flash memory cell to program (exploiting (a) FN tunneling or (b) CHEI) and to erase (exploiting (c) FN tunneling).

Fig. 1.3 shows the voltages schemes required to exploit both physical mechanisms to program or erase a Flash memory cell. To achieve FN tunneling, a large enough vertical field must be developed in the tunneling oxide. Source and drain contacts are left floating or at the same potential of the substrate to avoid the breakdown of the body-to-source and body-to-drain junctions, V_{CG} is either positive or negative with respect to the body potential (to perform a program or erase operation, as shown in Fig. 1.3(a) and 1.3(c), respectively). On the other hand, a combination of horizontal and vertical electric fields to accelerate the electrons and redirect them toward the FG is needed for CHEI, therefore the source and body contacts are kept to ground, while V_{CG} and V_D are moved to positive (but different) voltages (see Fig. 1.3(b)). In Fig.1.3 the erase operation is shown, which must be necessarily performed by FN tunneling (that is because it is impossibile to accelerate electrons stored in FG).

The two different mechanisms have preferred target application in which they perform better than the other. One of the main differences is that FN tunneling requires much higher voltages than CHEI to obtain the same ΔV_T over a comparable time scale [8]. On the other hand, during CHEI only a small fraction of the charge flowing from drain to source is redirected toward the FG contributing to the ΔV_T, while during a FN program operation almost all the current flowing is injected. If we define the injection efficiency as the ratio between the current injected into the storage layer and that flowing through the remaining contacts, it is very close to 1 for FN while it drops to $10^{-6} \div 10^{-7}$ for CHEI [5, 9].

1.2.3 Memory arrays

The device introduced in the previous section can be organized and operated in different ways depending on the target application of the resulting memory technology. The NVSM market is dominated by the NAND and NOR Flash technolo-

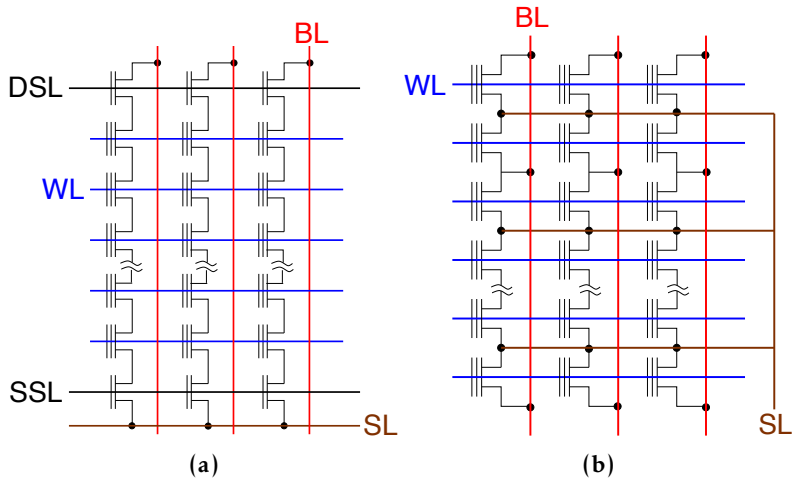


Figure 1.4: Schematic of the FG memory transistors arranged according to the (b) NOR Flash and (a) NAND Flash memory architecture.

gies. They are both based on the aforementioned memory transistor but they differ in the design of the single memory cell, in their array structure and operating conditions. However, as will be explained in the following section, the main feature they share is the erase operation being performed simultaneously on a large number of memory cells (called respectively sector or block for NOR and NAND Flash arrays), in a Flash (hence their name).

The architectures of both a NOR Flash and NAND Flash array are shown in Fig. 1.4. NAND Flash architecture is designed to minimize the chip cost per bit (the ratio between the capacity of the chip and its area), while maximizing, at the same time, the operating data throughput, as the technology is meant to store large amount of data. To follow these requirements, the number of contacts and service elements to address each memory cells of the array and the segmentation of the array have been constantly reduced over the years. Therefore, as shown in Fig. 1.4(a), in a NAND Flash array the memory transistors (up to 128 in state-of-

the-art technologies) are connected in series into strings that share the same horizontal *word-lines* (WLs), without requiring any area-consuming contacts at the drain or source side of the memory. To connect the strings end to the *bit-lines* (BLs, running vertically parallel to each string) and the *source-line* (SL, in common to all the strings) two select transistors at the top and the bottom of each string (connected to the corresponding source-select-line, SSL, and drain-select-line, DSL) are employed. The main drawback is that long times are needed to access randomly selected bits in the array. NOR Flash technologies are designed to meet the specifications of code-storage application, and they aim to provide fast access to smaller packets of data. A stronger array segmentation is employed to reduce the time delays coming from long signal lines and, within each array section, as shown in Fig. 1.4(b), the memory cells are connected in parallel. The result of such parallel connection is that both the gate and drain contacts of each cell are available, enabling a fast and single-cell selective program operation by adopting CHEI. The NOR Flash architecture drawback is the larger number of contacts needed (with respect to NAND Flash), that increasese the cell footprint ($\approx 10F^2$, twice the size of a NAND Flash [8]).

1.3 Scaling, reliability and novel design

50 years ago Moore predicted a reduction of devices feature size by a $\sqrt{2}$ factor every 2 years [8, 10], the so-called Moore's law. Flash technologies have been following the rule, bringing several advantages: the scaling process allowed to operate Flash arrays with lower voltages and reduced power consumption together with an increase of the bit storage density and a consequent reduction of cost. However, the pursue of this aggresive scaling has been stopped due to some fundamental problems related not only to the practical fabrication process but also inherent limitations of the physical

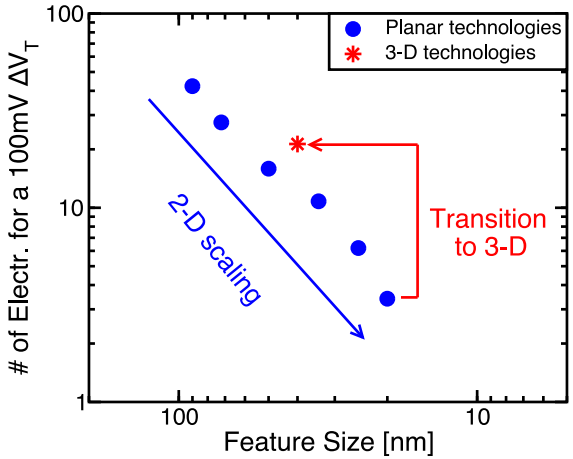


Figure 1.5: Number of electrons needed to achieve a V_T-shift equal to 100 mV for different NAND Flash technology nodes (from [10]). The red arrow describes the evolution toward 3-D NAND Flash architectures, which will be presented in § 1.3.2 .

mechanisms involved in the device operation. The process related problems are connected to the commonly adopted argon/fluorine immersion lithography [11] that could not achieve feature sizes smaller than 40 nm. In order to surpass this limitation and reach half the feature size, double-patterning techniques were introduced [11] and, in principle, more advanced quadruple-patterning allows to reach the 10 nm technology node [11]. At any rate, due to the high cost and process complexity, the smallest (NAND Flash) commercialized memory has a feature size of 14 nm [12].

1.3.1 Reliability issues

As mentioned above, together with process limitations, also known physical limitation become more and more relevant with the shrinking of the device dimensions: Program Noise (PN) and Random Telegraph Noise (RTN).

PN is related to the variability of the programmed V_T, Eq. 1.1

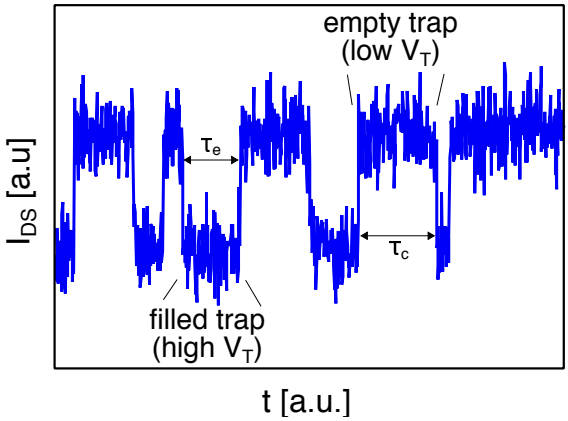


Figure 1.6: Plot illustrating the telegraphic behaviour of the I_{DS} fluctuations in time due to single-trap RTN in a Flash cell.

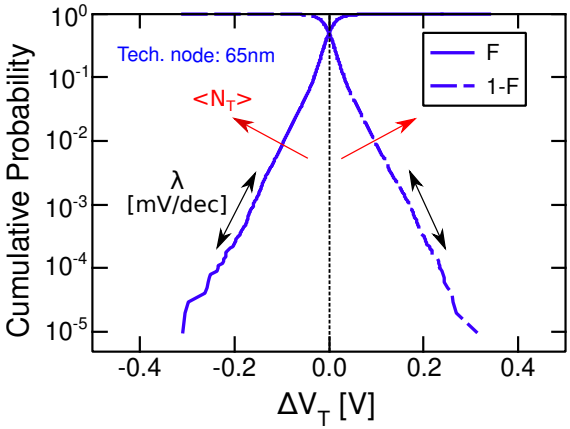


Figure 1.7: Cumulative distribution of ΔV_T between two consecutive measurements in time on a NOR Flash array (adapted from [13]).

shows that the ΔV_T of a memory cell is related to Q_s through C_{pp}. Therefore for small C_{pp} even small amounts of charge stored in the FG determine a non-negligible ΔV_T, i.e. if C_{pp} become as low as a few tens of aF, even a few electrons generate a non-negligible ΔV_T. The fact that the number of electrons injected to the FG from the substrate is non-deterministic and subject to uncertainity by nature, combined with the fact that even a few of them impact the V_T, give rise to fluctuactions of V_T [14, 15], which in turn give rise to uncertainty on the stored data.

RTN is a reliability issue that depends on traps localized in the tunnel oxide and at the interface between silicon and oxide. The presence of these traps can give rise to trapping and detrapping events which can result in abrupt variation of the cell V_T. This effect can be observed by measuring I_{DS} at constant bias voltages. Fig. 1.6 shows an example of such current fluctuactions between a low and high well defined current levels, along with the charateristic times of capture and emission, respectively τ_c and τ_e. This shows us that only one RTN trap is involved and that the current is high when the trap is filled and low when it is empty. Since the capture/release process is inherently random and the ΔV_T contribution from each trap is different from the others, the V_T instabilities in a Flash array due to RTN are described statistically [13]. The statistical distribution F of ΔV_T together with its complementary distribution 1-F, measured on the cells of a Flash array, are shown in Fig. 1.7. The strong signature of RTN fluctuactions can be seen in the low probability exponential tails [13]: the slope of these tails, λ (mV/dec) is the main parameter that defines RTN; the larger λ the more severe the impact of RTN, since larger V_T oscillations can be found in the memory arrays with higher probability. The deep scaling of modern memory cells bring an increase of λ and this increase is related to the average ΔV_T due to a single RTN trap [13]. Finally, together with the shrinking of

the cells dimension, also the WL and BL pitches are reduced during the scaling process, making the packing of Flash cells tighter and tighter. The capacitive coupling between consecutive cells is enhanced by the reduction of the distance between them, leading to reliability issues related to cell-to-cell interference. As a result, the injection of charge into the FG of one cell can induce an unwanted V_T shift also in the neighbouring ones, thus compromising the information stored in them [16].

1.3.2 Transition to 3D architectures

Although in principle device scaling represents the most straightforward path to improve Flash memory performances and reduce their cost, from what has been put forward above it becomes evident that some unwanted drawbacks are unavoidable. Since NAND Flash memories aim to high integration density, their improvement has been the real driving force behind the efforts to push the technology scaling to its physical limits, and NAND Flash memory cells have reached the smallest feasible feature size equal to 14 nm in the middle of 2010's decade [10]. To overcome such physical limits the elected solution was the exploitation of the third dimension, consisting in stacking many memory cells in the vertical direction, thus relaxing the requirement on the dimension of a single cell, while increasing, at the same time, the array integration density. Also the geometry of the memory transistor has been adapted to this change by adopting a cylindrical gate-all-around structure [10]. The result of this process is the birth of three dimensional (3-D) NAND Flash memory arrays, which have allowed to prolong the performance improving and cost reduction of non-volatile semiconductor memories for the storage of large amount of data.

The trend of the feature size of planar NAND technology and the number of stacked WL layers for 3-D NAND arrays during the last years are shown in Fig. 1.8(a). It can be seen

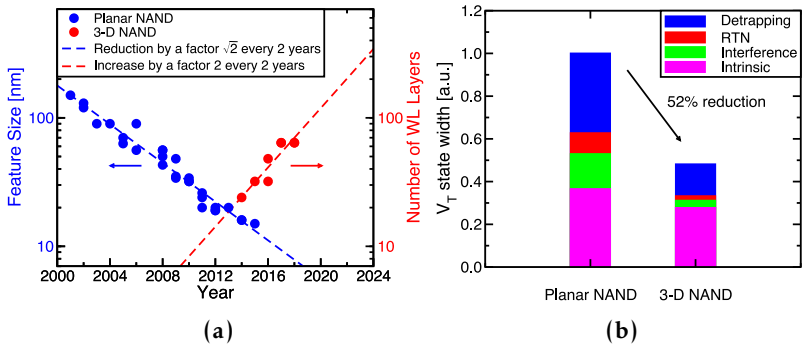


Figure 1.8: (a) Trend of the feature size of planar NAND technology and the number of stacked WL layers for 3-D NAND arrays during the last years [17]. (b) Relative impact of the variability issues discussed in § 1.3.1 on the total V_T distribution width for a 20 nm planar NAND and 3-D NAND Flash technologies [18]; moving from the former to the latter results in more than 50 % improvement.

that from 2001 to 2015 the industry pursued the increase of storage density by reducing the single cell dimension, resulting in halving the device area every two years according to Moore's law, while, starting from 2014, the trade-off between cell area and integration density has been broken by the introduction of 3-D NAND Flash technology, since cell miniaturization is now replaced by stacking many memory layers in the vertical direction. Thanks to the possibility of keeping the memory cell dimension larger and thanks to its geometrical structure, the transition from planar to 3-D technologies has relieved the reliability issues mentioned in § 1.3.1 (program noise, RTN, detrapping and cell-to-cell interference). The relative impact of these issues on the width of the V_T distribution is reduced by more than 50 % as shown in Fig. 1.8(b).

1.4 Volatile memories

As already mentioned in § 1.1, volatile memories cannot retain data when power is switched off, but they are much quicker

in writing and reading data compared to non-volatile memories. The most widespread general purpose volatile memory is the so called random-access memory (RAM).

There are two different kinds of RAM:

- **Dynamic RAM**: as shown in Fig. 1.1(b) is the most widely used RAM due to its cost-effectiveness. It is commonly called DRAM and employs one capacitor and one transistor to store one bit. It will be described in more detail in § 1.4.1.
- **Static RAM**: usually called SRAM, is much faster than DRAM but needs six transistors to store one bit and this is reflected in the very high costs that relegate it to niche uses.

1.4.1 DRAM cell operations

As shown in Fig. 1.9, the DRAM cell is composed by one transistor connected in series to one capacitor. The data are stored via the charge on the capacitor's plates, which is therefore called storage capacitor, while the MOSFET is called access transistor because it is used to select the capacitor on which the read/write operation has to be performed [19].

We define the supply voltage V_{DD} as the logic level '1' and ground the logic level '0'. During the write operation the bit-line is biased at the required level (1 or 0) and once the access transistor is powered-up the bias is transferred to the storage capacitor. During the read operation the bit-line is first biased at $V_{DD}/2$, and as soon as the MOSFET is powered up the stored charge is shared between the capacitor and the bit-line capacity. The bit-line bias will increase or decrease depending on the presence or absence of charge stored in the capacitor, the resulting signal being written as:

$$V_{signal} = \frac{1}{2} \cdot V_{storage} \cdot \frac{C_{storage}}{C_{storage} + C_{bitline}} \tag{1.2}$$

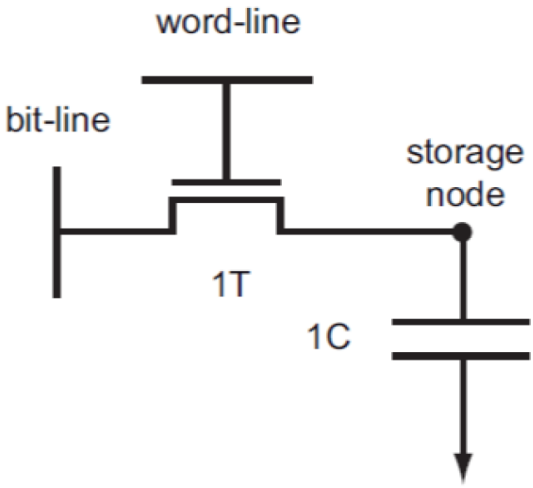
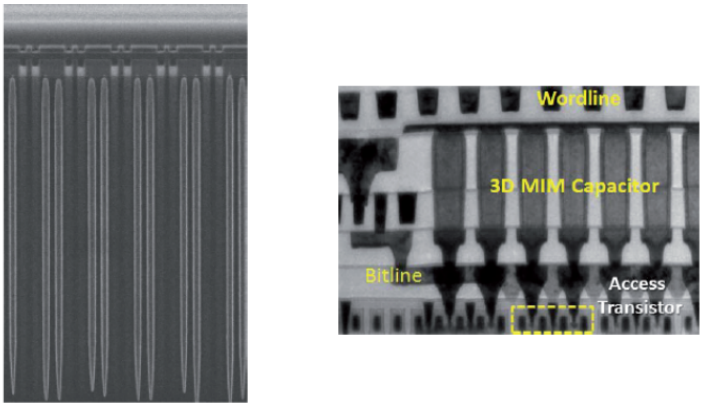


Figure 1.9: DRAM cell scheme.

where $V_{storage}$ is the bias difference between 0 and 1 on the storage capacitor and $C_{bitline}$ is the parasitic capacitance of the bit-line which includes the input capacitance of the sense amplifier that is employed to actually read the bias. It is worth saying that the read operation is destructive, i.e. the data has to be rewritten again (refresh) because the charge on the capacitor is lost once transferred. From a physical point of view, the implementation of the storage capacitor is one major challenge: as can be seen from Eq. 1.2 the storage capacitance must be as high as possible. The classical solution employs a polysilicon layer that must be deposited with the substrate. This solution has the major benefit of not requiring any additional steps with respect to the plain CMOS process but also the disadvantage of being very area consuming.

1.4.2 DRAM scaling

Since their first appearance, as the Flash memory cell, DRAM cells has been subjected to an intensive scaling process with



(a) (b)

Figure 1.10: SEM images of 22 nm DRAM cell with (a) deep-trench [20] and (b) stacked [21] storage capacitor.

the same aim of increasing the integration density to reduce the cost per stored bit. During the scaling process from 20 nm to 10 nm in the storage capacitor implementation, new problems and challenges arose. The value of $C_{storage}$ is the most critical parameter that strongly impacts the V_{signal} that has to be amplified, the speed of the DRAM cell and the data retention time. It is usually accepted that the capacitance value has to be at least 25 fF [22]. Maintaining this value with the aggressive cell scaling posed serious challenges that have been answered with new geometric structures and new materials. In particular, it has been possible to maintain such capacitance values employing three dimensional structures deposited above the silicon surface (stacked capacitors) or within the silicon substrate (trench capacitor) as shown in Fig. 1.10 [20] [21]. The trench capacitor is implemented by the etching of holes in the substrate filled with a layer of oxide sandwiched between two polysilicon layers that acts as contacts. The stacked capacitor is implemented above the access transistor by stacking a few layers of polysilicon. The aspect ratio of these structural solutions has nowadays reached values

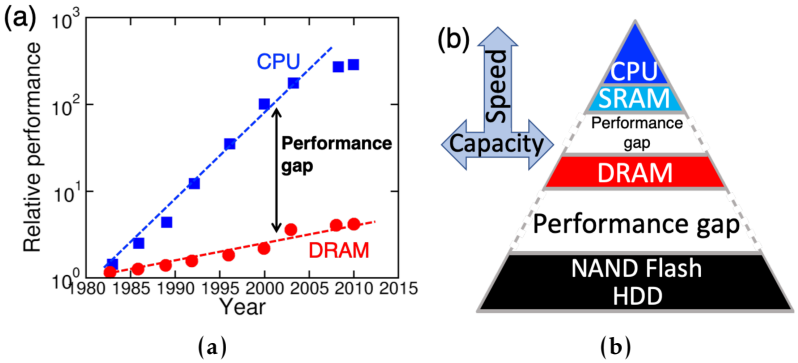
higher than 50. Alongside the implementation of a three dimensional structure, in order to enhance the capacitor area, while mantaining the same planar area consumption, research efforts have been directed toward new materials with high values of dielectric constant (high-k) to be placed inside the capacitor. In order to satisfy the roadmap targets [23], i.e. equivalent oxide thickness values less than 0.5 nm and leakage currents in the order of 10^{-8} A/cm^2 when biased at 1V, it is needed an oxide with k> 50. Nevertheless, because they usually have a narrow bandgap and thus high leakage current, these high-k materials present a trade off. Another scaling related issue is poor retention of written data due to the increasing leakage currents that discharges the storage capacitor. To enhance the transistor performances, the reduction of the channel length has been compensated by an higher substrate doping and a lower oxide thickness. These two counter measures actually worsen the leakage currents: the former allows bigger currents through the capacitor junction, while the latter allows higher tunneling currents toward the gate. To avoid data loss, a periodic refresh of the DRAM cell is needed, which usually takes place every 64 ms [24]. Reducing this period might be a viable solution, but it comes with the cost of higher power consumption. All the aforementioned issues related to the sub-20 nm, directed the research toward novel structures to implement DRAM cells. Since the 3D capacitor aspect ratio reached extreme levels (>50), a viable solution was manufacturing a DRAM cell without the storage capacitor, the so-called *capacitorless* DRAM [25]. One of the most common versions of this type of cell simply uses a transistor with an isolated and floating substrate, in which holes accumulations are created by means of appropriate bias conditions at the contacts of the device. These holes build-ups alter the threshold voltage (V_T) of the transistor, storing the data. At any rate, this cell shows relevant reliability issues: V_T instabilities, short channel effects, high leakage currents in

the off state and non-uniform doping effects. In order to overcome these issues, new schemes and structures have been formulated in the last years, such as: ultra-thin SOI MOSFET, Z-RAM and vertical double-gate single-transistor cell [26]. Solutions that exploit different physical mechanisms have also been explored, like the thyristor cell [27]. However, none of these solutions has proved sufficiently advantageous to replace the basic structure of the 1T-1C cell.

1.5 Von Neumann-bottleneck

The previous chapters presented the volatile and non-volatile memories dominating nowadays the market, together with the main issues the scaling of these technologies has faced, and innovations that have partially allowed to follow the previous scaling trend. In addition to the scaling issues mentioned above, today's information technology has to face another major challenge, the so-called *von Neumann bottleck*. The great majority of computer systems is based on the von Neumann architecture, which is characterized by a rigid separation of logic and memory circuits [28], i.e. the processing unit (CPU) and the working memory (typically a DRAM) are physically separated.

Fig. 1.11(a) shows the performance exponential growth of both central processing unit (CPU) and DRAM [19] [30]. From the graph, it is clearly visible the presence of a performance gap in term of cycle time that is getting bigger and bigger since forty years ago [29]. This slower improvement in cycle time has produced a memory bottleneck that could reduce a computing system overall performance and it is called memory wall or [23] von Neumann bottleneck [31]. To mitigate such gap, a caching layer of SRAM has been introduced between the CPU and the DRAM employed for the working memory [32] [33]. Since both the DRAM and SRAM are volatile memory technologies, a non-volatile memory, such as the afore-



(a) (b)

Figure 1.11: a) Relative performance evolution for CPU and DRAM. Even though they share the exponential growth dictated by Moore's Law, an increasing performance gap formed over the years. Adapted from [29]. b) Schematic representantion of performace hierarchy.

mentioned NAND Flash, must be introduced to ensure the data retain. A schematic representation of the memory hierarchy in a computer system is shown in Fig. 1.11(b) [34]. It is worth noting that there is another performance gap between the DRAM layer and the NAND Flash one.

1.5.1 Emerging memory technologies

One of the proposed solutions to mitigate the von Neumann bottleneck employing an improved storage is the Storage Class Memory by IBM. The idea is to fill the performance gap between the working memory (RAM) and the storage memory (usually NAND Flash). In order to fill this gap, new technologies should have high read/write speeds, below 100 ns like DRAM, low cost per bit, high density and non-volatility like Flash memory. There are many promising candidates for SCM implementation [35] and more in general to overcome some of the issues presented throughout this chapter, among them we find (non-exhaustive list):

- **Resistive RAM (RRAM):** consists of a metal-insulator-

metal (MIM) structure where a dielectric layer is sandwiched between two metal contacts. A localized defect-rich path, called conductive filament (CF), can be induced and destroyed, in the dielectric layer, through suitable voltage pulses and the presence or absence of such CF can change the resistance of the dielectric up to 5 orders of magnitude. They are very fast and can work with currents as low as $15\mu A$ but they show high variability that result in reliability issues.

- **Phase change memory (PCM):** also consist of MIM structure where the insulator material is a chalcogenide phase-change material, i.e. a material that exists in two different solid states: amorphous and crystalline. The switch between the two phases is triggered by temperature, but exploiting the Joule heating can be induced with suitable voltage pulses. Despite the bulk-type switching, as opposed to filamentary switching in RRAM, the PCM also suffers from variability effects, which are critical in hindering memory applications. Nevertheless PCM are considered promising due to their sub-100 ns switching time and more than 10^9 cycling endurance [36].

- **Magnetoresistive random access memory (MRAM):** the fundamental building block of an MRAM is the magnetic tunnel junction (MTJ), which consists of an MIM stack where the conductive electrodes are made of a ferromagnetic material and the insulator has a thickness of 1 nm. The magnetic polarization is fixed on one of the two contacts while can be switched on the other. The magnetic polarization of the two electrodes can be parallel or antiparallel and thanks to the magnetoresistive effect [37] the tunneling probability and thus the resistance of the junction can changed by such magnetic polarization switching. The magnetic polarization can be switched by applying a magnetic field such in Toggle-

MRAM [38] or by the spin-trasnfer torque (STT) effect. [39] STT-MRAM is attracting a strong interest due to its fast switching [40], high endurance [41], CMOS compatibility and low current operations [42].

- **Ferroelectric memories:** are a class of memory technologies which will be presented in the next chapter, with particular focus on ferroelectric tunnel junctions (FTJ), which is the main topic of this thesis. Ferroelectric memories exploit the ferroelectricity of a material, i.e. the ability to have a stable electric polarization even in the absence of an external polarizing electric field. The interest in this type of memory technologies has increased abruptly with the discovery of the ferroelectricity of hafnium oxide (HfO_2) which is a CMOS compaible material.

Chapter 2

Ferroelectric memories

This chapter begins with a brief introduction on ferroelectricity, explaining its origin and its relevance as an exploitable phenomenon in memory technology. A non-exhaustive list of ferroelectric materials is also provided. Subsequently, an overview on the so far most successful ferroeletric memories shall be presented: ferroelectric RAM (FeRAM), ferroelectric transistor (FeFET) and ferroelectric tunnel junction (FTJ). FeRAM is a modified DRAM cell in which the capacitor oxide is substituted with a ferroelectric material. FeFET, similarly, is a memory transistor in which the gate oxide has been replaced with a ferroelectric material. FTJ structures shall be described in more detail: from the concept proposal in 1971 to the MIM structure and metal-insulator-semiconductor (MIS) structure to more advanced concept employing buffer layers to create crested barriers. A detailed mathematical physical description of FTJs shall also be given. At the end of the chapter an overview on the recent FTJ concepts present in literature shall be presented.

2.1 Polarization

When an insulator material is embedded in an electric field E [V/m] the positive and negative charges inside the material are separated creating dipols. This phenomenon is called polarization and it results in a superficial macroscopical charge, described through the polarization vector **P** [C/m^2]. The relation between E and P is described by the following equation:

$$P = \epsilon_0 \chi E \tag{2.1}$$

Where ϵ_0 is the vacuum electric permittivity and χ is the electric susceptability of the material, which is related to the relative permittivity ϵ_r through the relation:

$$\chi = \epsilon_r - 1 \tag{2.2}$$

2.1.1 P-E relations

As shown in Fig. 2.1, there are four different relations between P and E that characterize insulator materials:

- **a) dielectric:** χ is constant and the relation is linear. Although it is not an accurate definition, insulators are often called dielectrics, due to the fact that this relation is the most common one.
- **b) paraelectric:** χ is not constant but it is field-dependent. And, usually, the relation becomes linear again at high fields.
- **c) ferroelectric:** not only the relation between field and χ is non-linear but the P-E relation presents a hysteretic behavior, i.e. the value of the vector P depends not only on the current value of the electric field but also on its history. As a resulting effect, one can clearly see the presence of two states with polarization different from 0 in the absence of an external polarizing electric field. A

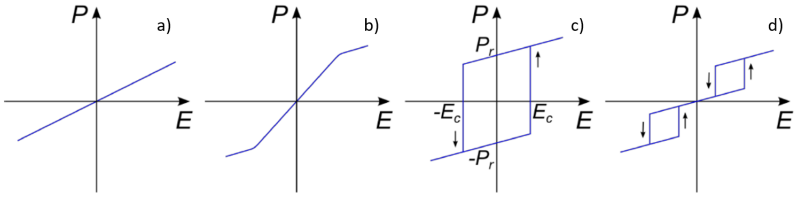


Figure 2.1: Polarization-electric field relation for a) dielectric, b) paraelectric, c) ferroelectric and d) antiferroelectric.

more detailed mathematical description will be given in § 2.1.2.

- **d) antiferroelectric:** hysteresis is again present but only with sufficiently high absolute value of the electric field. As a resulting effect the stable non-zero polarization states are absent.

Thus, a ferroelectric material can expose a permanent electrical polarization even in the absence of an external polarizing electric field. There are two stable non-zero polarization states and the polarization vector can be switched between the two states by means of an high enough electric field. This phenomenon will be better described in the next section.

2.1.2 Ferroelectricity

Ferroelectricity was first discovered in 1920 by the PhD student Joseph Valasek (1897–1993), who was working under the supervision of William Swann at the University of Minnesota, Minneapolis, US. From a reticular point of view, only non-centrosymmetric structures can lead to ferroelectricity, where centrosymmetric structures are the ones that possess an inversion center. Furthermore, all ferroelectric materials are themselves pyroelectric and piezoelectric. The former is a phenomenon in which polarization is related to the temperature of the material, while the latter is a relation between

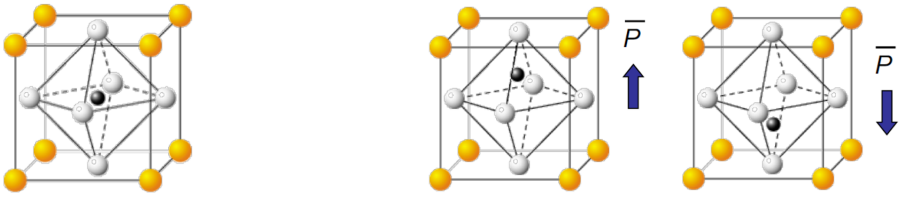


Figure 2.2: (a) Ferroeletric material at $T > T_C$ recover a centrosymmetric structure and loses ferroelectric properties and (b) ferroelectric material for $T < T_C$.

polarization and mechanical stress; non-centrosymmetry is a prerequisite also for this two phenomena.

All ferroelectric materials present a characteristic temperature named **Curie temperature**, along the lines of the characteristic temperature of ferromagnets, above which the centrosymmetric structure is recovered and all the ferroelectric properties are lost (Fig. 2.2(a)). Below this temperature the centrosymmetric structure becomes unstable, while the two different deformed cells shown in Fig. 2.2(b) do not. A mathematical description within the framework of Landau phase transition theory [43] will be given in § 4. The presence of piezoelectric and pyroelectric effect in ferroelectric materials means that temperature and mechanical stress strongly affect the polarization of the material and thus obtaining stable polarization in a nanometric memory device such as FTJ is not an easy task. To describe the hysteresis cycle a few significant biases and polarization are defined, as indicated in Fig. 2.3:

- **Saturation polarization P_s:** is defined as the maximum polarization of the hysteresis cycle. All electric fields above this point will produce a linear, dielectric-like polarization.
- **Remanent polarization P_r:** is defined as the absolute value of the polarization when the electric field is 0. Note that

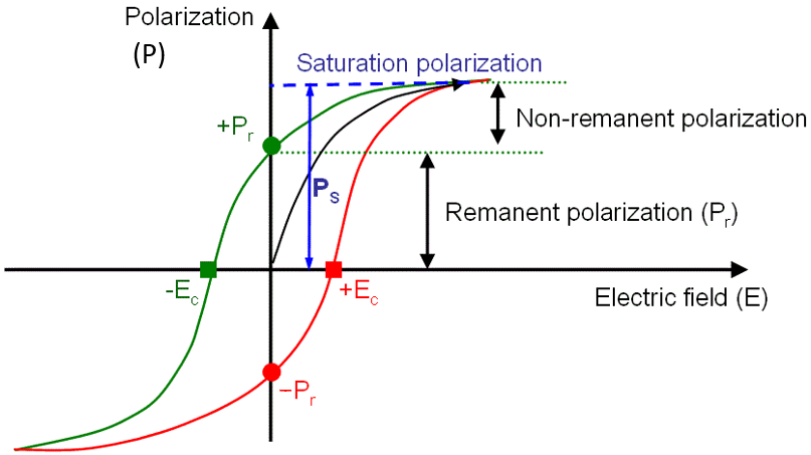


Figure 2.3: Hysteresis cycle diagram in which significant biases and polarization values are highlighted.

the two stable states have equal and opposite polarization.

- **Coercive field E_C:** is defined as the absolute value of the field that is able to switch the polarization to the other direction, i.e. if we are at the point in which P is equal to $+P_r$ and we apply $-E_C$ we will end with a polarization vector of $-P_r$, even if we switch off the electric field.

2.1.3 Ferroelectric materials

Perovskite crystals are the most investigated and widespread ferroelectrics. They have the general formula ABO_3, in which the cation A has valence number between +1 and +3, whereas the cation B has it between +3 and +6. The most employed materials in this group are lead-zirconate titanate (PZT) and barium titanate (BTO). Fig. 2.4 shows a perovskite unit cell in the ferroelectric state, i.e. a non-centrosymmetric structure. The ferroelectricity of the material derives from the possibility of certain atoms in the unit cell to reside in two

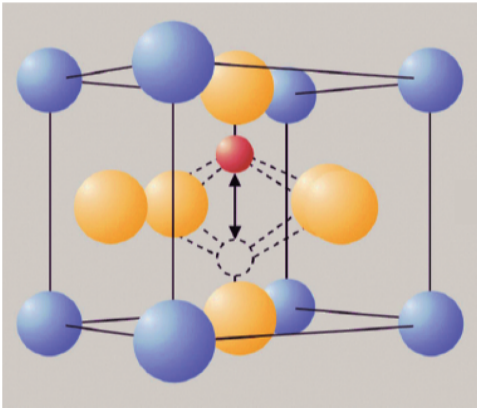


Figure 2.4: Schematic illustration of a perovskite unit cell with ABO_3 structure. For PZT we have A=Pb (blue), B=Zr or Ti (red) and the oxygen atoms O_3 (yellow).

different stable energy positions. This causes a shift in the charge centre of the cell and thus results in a permanent polarization of the material, even in the absence of an external electric field. However, by applying a sufficiently high electric field, the aforementioned coercive field, it is possible to switch from one to the other stable condition. Ferroelectric materials based on perovskites have been intensively studied for memory applications, but their non-compatibility with the CMOS process has limited their adoption. Ferroelectric properties in oxides with crystallographic structure like fluorite, as hafnium oxide HfO_2 and zirconium oxide ZrO_2, have been reported for the first time in 2011 [44] and generated a lot of interest toward ferroelectric memories since hafnium oxide has been employed as high-k material (see §1.4.1) in advanced CMOS processes. In literature it is reported that, in order to achieve ferroelectric behavior in HfO_2, it must be doped with 3 – 10 mol% of Si,Y,Al,Gd or Sr [45] and have it crystallize in films of thicknesses in the order of 10 nm. The discovery of a CMOS-compatible ferroelectric material opened the possibility of overcoming the limits of the known

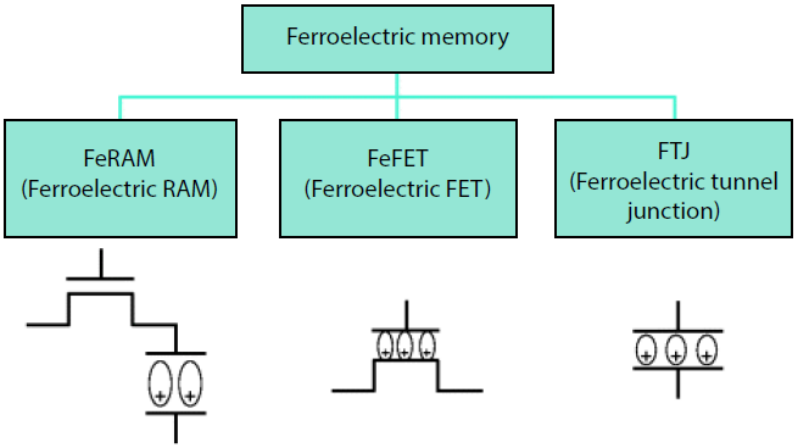


Figure 2.5: Schematic illustration of the ferroelectric memory types.

ferroelectrics in their application to memory technologies.

2.2 Ferroelectric memories

As shown in Fig. 2.5, there are three main types of ferroelectric memories: FeRAM is based on the DRAM 1T-1C cell where the oxide material is a ferroelectric one, FeFET exploits the same idea and the ferroelectric material replaces the gate stack, FTJ presents a nanometric layer of ferroelectric material sandwiched between two conductive electrodes. Every technology has been studied with different target applications in mind: FeRAM and FTJ are considered promising solutions for the post-DRAM scenario, while FeFET has the aim of replacing NAND Flash cells.

2.2.1 FeRAM

FeRAM is an evolution of a DRAM 1T-1C cell where the capacitor insulator is replaced by a ferroelectric material. In a standard DRAM cell, information is kept through the charge stored in the capacitor, while in a FeRAM is kept through

Figure 2.6: (a) TEM imaging and (b) P-E relation of a 3-D capacitor TiN/HfO^2/TiN. A and B are the dimensions of the projected area which can be computed as A·B [46].

the remanent polarization P_r of the ferroelectric material. In this way it is possible to overcome the issues related to the leakage currents that discharge the capacitor and force a continuous refresh of the stored data; in a FeRAM the data is retained even without external supply. During the writing operation the access transistor is powered up and the bitline bias allows to force a certain polarization status in the ferroelectric layer and thus to store the data. The read operation is destructive also in this case since it is performed by applying a field greater than the coercive one and by sensing the magnitude of the current to understand if the polarization vector switched direction and thus the initial state of the cell. Memory matrices based on perovskite materials have been developed and commercialized without a great success, since their costs and performances could not rival the other technologies. Since to ensure a robust enough ferroelectricity a thickness of 70 nm is needed and the capacitors cannot be built with a 3-D design as in the standard DRAM case, continuous scaling of the device is impossible to obtain. In the planar case of perovskite-based FeRAMs, such as PZT or SBT, the main problem for scaling is the limited amount of charge for proper cell operation. According to Maruyama et

al. [47], a 90 nm technology node FeRAM cell requires $2P_r > 84\ \mu C/cm^2$ but for thin layer of crystalline PZT, $2P_r$ is limited to $< 80\ \mu C/cm^2$ and other materials have even lower values. Fabricating a nanometric 3-D ferroelectric capacitor might be a likely solution to these issues; yet, depositing perovskite structure with high aspect-ratio through *atomic layer deposition* could be rather challenging. Due to these difficulties, traditional ferroelectrics cannot be employed for technology node smaller than 130 nm. Ferroelectric layers based on HfO_2 may be a turning point for the implementation of 3-D capacitors. Polakowski et al. [48] report ferroelectric properties in the 3-D TiN/Al:HfO_2/TiN capacitor with a P_r per projected area of $152\mu C/cm^2$. In Fig. 2.6(a) a TEM image of a deep-trench capacitor with Al:HfO_2 and the respective P-E curve are shown. Although the P_r value for the planar capacitor was only $15\mu C/cm^2$, the 1.6-μm-deep-trench capacitor with 13:1 aspect ratio capacitor can increase the P_r per projected area by an order of magnitude, which is very similar to the increase in the capacitor area.

2.2.2 FeFET

The second type of ferroelectric memory is the FeFET. The idea behind the FeFET memory cell is to exploit the charges induced by the polarization of the ferroelectric layer to modulate the conductivity of a transistor channel. To achieve this, a ferroelectric material is inserted in place of the gate oxide of the transistor itself. Once data have been stored, they can be read non-destructively by measuring the drain current as a function of the gate voltage, i.e. via the trans-characteristic. Several FeFET structures have been considered. The first is the MFS (metal-ferroelectric-silicon) and consists of replacing the gate oxide of a standard MOSFET with a ferroelectric. Although being the simplest structure, it bears several disadvantages. Classical ferroelectrics such as PZT or SBT are unstable when deposited on a silicon substrate, due to the diffu-

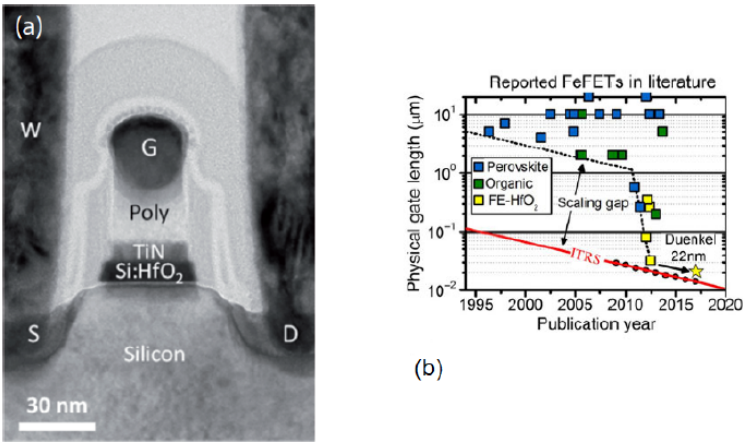


Figure 2.7: a) TEM image of a MFIS-FET with 10 nm of Si:HfO_2 and a SiON buffer layer of 1.2 nm. b) Physical gate length of FETs and FeFETs in literature [46].

sion of elements such as bismuth or lead in the substrate. Furthermore, the density of interfacial traps between perovskite and substrate is high enough to accelerate the degradation of the ferroelectric. The second structure, shown in Fig. 2.7(a), is the MFIS (metal-ferroelectric-insulator-silicon) structure, in which a buffer layer is inserted between the ferroelectric and the silicon. While counterbalancing the disadvantages of the MFS structure, this solution however brings new issues. Due to the low dielectric constant of the buffer layer, a strong depolarization field is created, which limits the retention of the device. Various solutions to solve this problem have been proposed, such as the insertion of a high-k buffer layer below the ferroelectric layer. The best retention properties for FeFETs with conventional ferroelectrics have been reported for cells formed by Pt/SBT/HfO_2/Si. As shown in Fig. 2.7(b), the technology gap of about two orders of magnitude that existed between logical FETs and conventional ferroelectrics before the adoption of ferroelectrics such as doped HfO_2 is now eliminated. In fact, to date, the smallest technological node re-

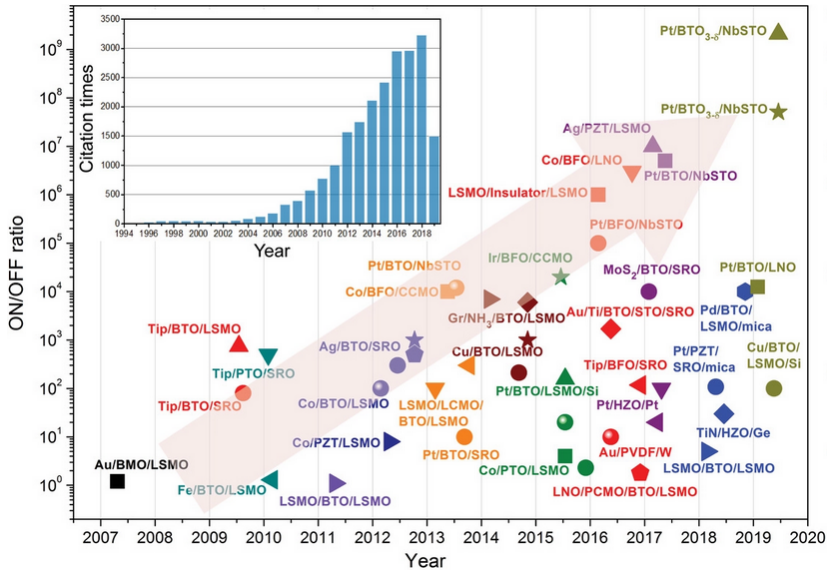


Figure 2.8: Relation between the year and the ON/OFF ratio achieved by various FTJ structures. The inset is citation times for FTJ devices per year, analyzed by Web of Science using the keyword "ferroelectric tunnel junctions", from [49].

ported for FeFET is 22 nm. In addition, the lower dielectric constant of doped HfO_2 compared to that of perovskites leads to lower depolarization fields and thus to an improvement in retention. Although the recent developments in the last decade, FeFET devices based on HfO_2 have retention characteristics comparable to or even better than those reported for the best perovskite-based cells. The high scalability and better retention have made doped HfO_2 the best oxide layer for non-volatile FeFET memories.

2.2.3 FTJ

The ferroelectric tunnel junction idea is having a ferrolectric barrier through which the current can pass by quantum tunneling effect and that the transparency of this barrier can be

varied through the direction of the polarization vector. Depending on the direction of the polarization vector, the device has two states: one characterized by low resistance ad high current called **ON** state or **LRS** and its opposite, with high resistance and low current, called **OFF** state or **HRS**. This concept was put forward for the first time 50 years ago, in 1971, by Esaki [50], but was then quickly disregarded, since it was believed that ferroelectricity could not be stable at nanometric thicknesses that are needed for quantum tunneling to take place. As shown in the inset of Fig. 2.8, the number of citation time for the keyword "FTJ" has been steadily increasing in the last decade and skyrocketed since the discovery of the ferroelectricity of the doped hafnium oxide. Fig. 2.8 shows the evolution of the technology in the last decade. The **ON/OFF ratio** is the figure of merit of the device and has been improved by nine orders of magnitudes in the last decade. It is defined as: ON/OFF ratio= I_{LRS}/I_{HRS}. However, this type of memory has a number of critical problems that cannot be easily overcome. During the manufacturing process, it is very challenging to achieve uniform thickness over the entire area of the device. Since the FTJ bases conduction on the tunnelling mechanism, even a small inhomogeneity on the atomic scale can have a significant impact on its operation. In addition, precise control of the interface between ferroelectric and electrode is required for optimum performance. To simplify the device design and relax the requirements on the ferroelectric layer, an alternative structure has been proposed with the addition of a thin dielectric layer [51]. In this structure, polarization switching takes place in the ferroelectric layer, but the tunneling current is limited by the dielectric layer. Although this device type and the proposed solutions are encouraging, many experimental and theoretical studies are still needed to bring the technology to the appropriate maturity level. In addition to this structure with added dielectric (often called *crested barrier FTJ*), as can be seen from Fig. 2.8, the basic

FTJ structure composed by a ferroelectric layer between two metal contacts has been modified during its development into Metal-Ferroelectric-Semiconductor structures (MFS). The advantages of such a structure will be explained in § 2.3.

2.3 FTJ: equations and performances

In the previous section the FTJ figure of merit was defined as the ratio between the tunneling current passing through the device in the low resistance state and the high resistance one. The tunneling current depends exponentially on the mean barrier height between the two contacts [52]. Therefore, to enhance the performances of an FTJ device we want to maximize the change in barrier height when the direction of the polarization vector is switched.

2.3.1 MFM structure

To understand the working principle of the device, an electrostatic analysis of the junction is needed. Since the ferroelectric material can expose a permanent polarization even in the absence of a polarizing field, a non-zero voltage drop in the ferroelectric can be present without any bias applied on the contacts. Such voltage drop is present also in the metal contacts if they are not considered as ideal metals. Following Thomas-Fermi theory [53], the potential profile inside a metal is defined as:

$$\phi(x) = \frac{\sigma}{\epsilon_r} \lambda e^{-\frac{x}{\lambda}} \tag{2.3}$$

where σ is the charge per unit area, ϵ_r is the relative dielectric constant of the metal and λ is the screening length of the metal. If we combine this equation with the Gauss law:

$$\sigma = -\epsilon_F F + \overrightarrow{P} \tag{2.4}$$

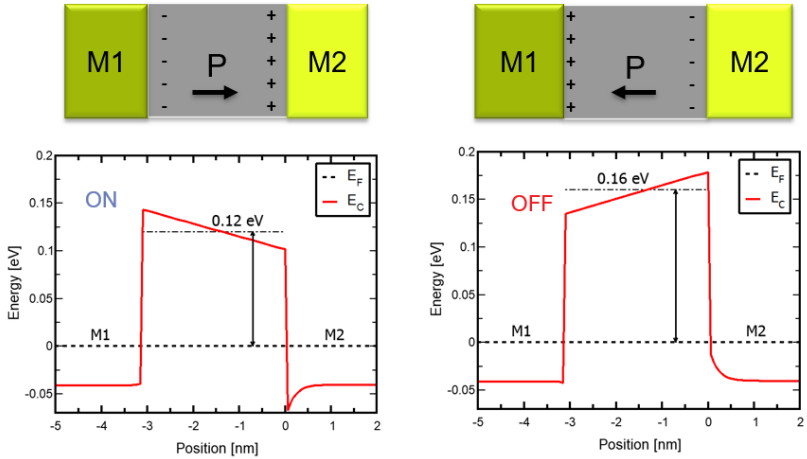


Figure 2.9: Schematic explanation of the working principle of an FTJ. On the left: polarization vector pointing toward the less ideal metal and lower mean barrier height, on the right is depicted the opposite case in which the mean barrier height is higher and thus the current is lower.

where F is the electric field in the ferroelectric layer, we can write the equation of the mean barrier height:

$$\overline{\phi_B} = \frac{1}{2}\left(\frac{\sigma\lambda_1}{\epsilon_1} - \frac{\sigma\lambda_2}{\epsilon_2}\right) = -\frac{1}{2}\left(\frac{\lambda_2}{\epsilon_2} - \frac{\lambda_1}{\epsilon_1}\right) \cdot \frac{Pd - \epsilon_F\phi_{bi}}{d + \epsilon_F\left(\frac{\lambda_2}{\epsilon_2} + \frac{\lambda_1}{\epsilon_1}\right)} \tag{2.5}$$

where the pedices 1 and 2 refers to the two metal contacts, ϕ_{bi} is the built-in potential of the junction, P is the magnitude of the polarization vector, d and ϵ_F are respectively the thickness and the dielectric constant of the ferroelectric layer. Finally, we can write the variation of the mean barrier height following polarization reversal:

$$\Delta\overline{\phi_B}^{(P)} = \frac{Pd\left(\frac{\lambda_2}{\epsilon_2} - \frac{\lambda_1}{\epsilon_1}\right)}{d + \epsilon_F\left(\frac{\lambda_2}{\epsilon_2} + \frac{\lambda_1}{\epsilon_1}\right)} \tag{2.6}$$

From Eq.2.6 it is evident that two different metal contacts

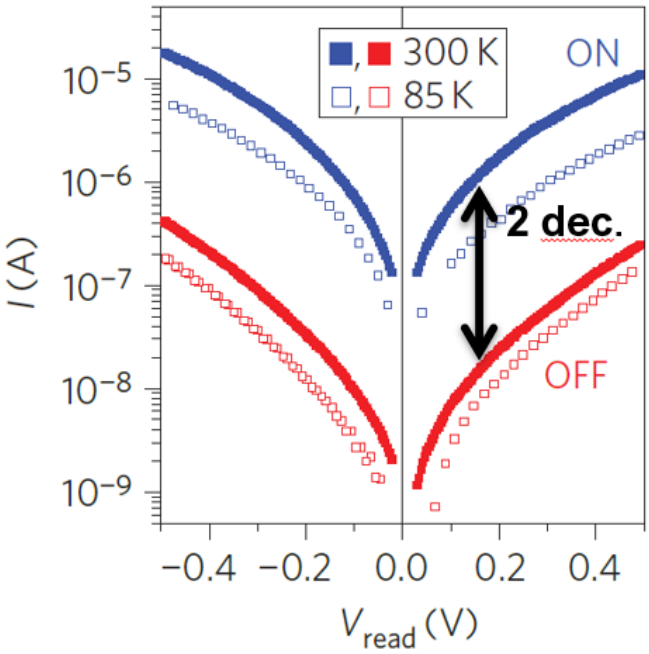


Figure 2.10: A resistive window of two decades in a MFM FTJ, from [54].

are needed to have a change in the mean barrier height upon polarization reversal and thus the existance of the LRS and HRS. Rearranging the same equations we can write the expression for the electric field in the ferroeletric layer as:

$$F = \frac{(\phi_{bi} + V) + P\left(\frac{\lambda_2}{\epsilon_2} + \frac{\lambda_1}{\epsilon_1}\right)}{d + \epsilon_F\left(\frac{\lambda_2}{\epsilon_2} + \frac{\lambda_1}{\epsilon_1}\right)} \tag{2.7}$$

where V is the bias applied to the junction. Note that even in the absence of built-in field and applied bias, the field in the ferroelectric layer is not zero. This field is called *depolarization field* since it works against the polarization vector, i.e. it tends to neutralize it. As can be seen from Eq. 2.7, the depolarization field is higher when the dielectric thickness d is lower and thus it plays a relevant role in FTJs, which

need nanometric ferroelectric layer by design. The described working principle is shown in Fig. 2.9: on the left we have the ON state, achieved when the polarization vector is pointing toward the less ideal metal (i.e. it shares the bigger part of the metal voltage drop); on the right, when the polarization vector is switched to the other side we have the OFF state characterized by a greater mean barrier height. As shown in Fig. 2.10, the tunneling current depends so strongly on the mean barrier height that even a change in the order of some tens of meV can result in dramatic changes in the current profile, here, the ON/OFF ratio, also called *resistive window* is of two orders of magnitude. As we will see in the next subsection, it is worth noting from Eq. 2.6 that the mean barrier height variation with the switching of polarization P is bigger if the screening length of the two metal are very different from each other. Considering that an ideal metal is modelled with a screening length equal to zero, it is almost straightforward to substitute the less ideal metal of the two, the one with the bigger λ, with an heavily doped semiconductor that can act almost as a metal contact while having a surely bigger screening length with respect to a real metal.

2.3.2 MFS structure

In the previous section we mentioned that the variation of the mean barrier height following polarization reversal, $\Delta\overline{\phi_B}^{(P)}$, is greater if the difference in the screening lengths of the contacts is higher. The screening length of a seminconductor is undoubtedly bigger than a metal one, but employing a heavily doped semiconductor as contact has another major advantage. As can be seen from Fig. 2.11, the working principle is similar to the one depicted in Fig. 2.9: when the polarization vector is pointing toward the semiconductor and exposes its positive charges at that interface, the mean barrier height is lower; in the opposite case, not only the mean barrier heigth is greater, but the presence of the negative charges of the fer-

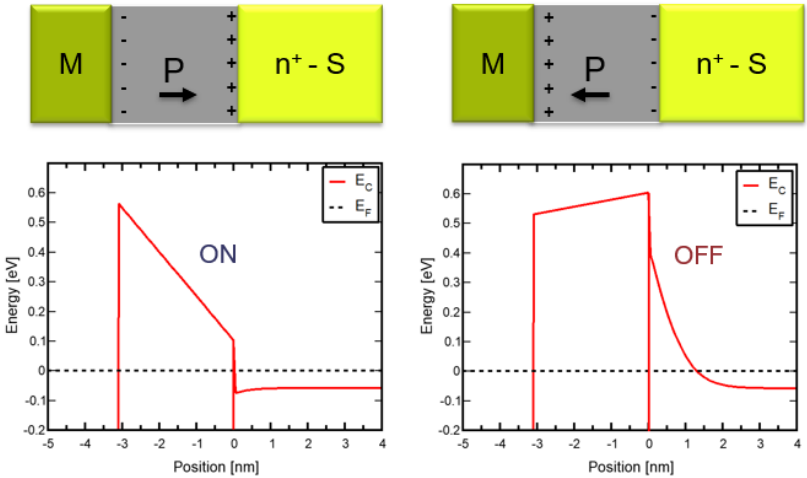


Figure 2.11: Schematic explanation of the working principle of an MFS FTJ. On the left: polarization vector pointing toward the less ideal contact (the semiconductor) and lower mean barrier height, on the right it is depicted the opposite case in which the mean barrier height is higher and its thickness is enhanced by the depletion layer of the semiconductor, and thus the current is lower.

roelectric remanent polarization at that interface makes the semiconductor depleted of their electrons and thus the resulting barrier is not only higher but also thicker. This effect strongly enhances the FTJ performances in terms of ON/OFF ratio. As can be seen from Fig. 2.12, the resistive window for an Pt/BTO/Nb:STO FTJ is doubled (four decades) with respect to the MFM structure. From Fig. 2.8 we can desume that the MFS structure has much better performances in term of ON/OFF ratio, even if it introduces other issues that are less severe with metal electrodes, such as the role of the oxygen vacancies migration across the interface [56] or the chance of Metal–Insulator phase transition to take place [57] or even the increased complexity of writing an analytical model that takes into account the semiconductor electrostatics and transport equations.

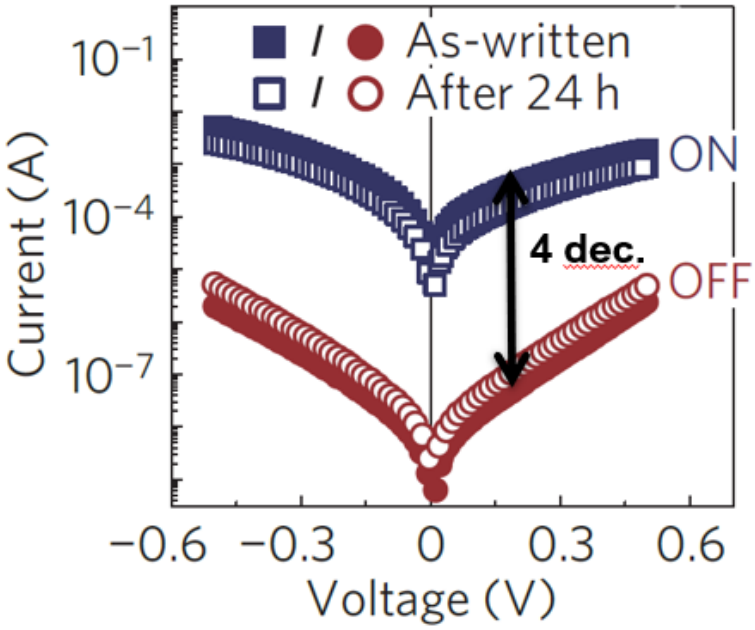


Figure 2.12: A resistive window of four decades in a MFS FTJ, from [55].

2.3.3 Purpose of the Ph.D. research

Even though MFS junctions have been extensively studied in the last decades, there are various issues that still need to be addressed. The research activities carried out during the Ph.D. attempt to contribute to fully understand FTJs working principles and how to improve their performance. The activities can be divided into two parts. During the first one, an experimental characterization of various Pt/BTO/Nb:STO FTJs has been performed, which confirmed the chance of achieving a large resistive window in such a device. Moreover, by measuring the current-voltage curve of the devices quickly after the set/reset pulses used for polarization reversal, the importance of taking into account the detrimental effects of the large depolarization fields in the ferroelectric layer when assessing the device memory window is demonstrated. The

collected experimental data are then reproduced with an in-house 1-D simulator written in MATLAB, which provided us with much more freedom with respect to commercial TCAD softwares; such as the chance of taking into account the tunneling current not only at the interface but also through the same conduction band of the depleted semiconductor (i.e. distributed tunneling). The model accounts also for drift/diffusion in the semiconductor and allows to catch a comprehensive physical picture of device operation. It represents a powerful tool for the design and optimization of ferroelectric tunnel junctions with semiconductor electrodes.

The second part focuses on modeling the switching process of the FTJ device. By coupling the aforementioned 1-D simulator with a novel in-house 2-D simulator describing the ferroelectric area in the Ginzburg-Landau-Devonshire theory framework, a self-consistent simulation of the device switching process can be obtained. Although being just preliminar ones, the results are pretty encouraging, and can be improved by means of further physical and electric characterization to guide the tuning of the custom model.

Chapter 3

Experimental characterization and 1-D modeling of current transport

This chapter presents the experimental and modeling activities on FTJ current transport. Throughout § 3.1 the experimental setup employed and the FTJ samples are presented. In § 3.2 the types of measurements, the relative schemes and the results are shown. The following subsections (from § 3.3) describes the 1-D model employed in the custom simulator and its implementation in Mathworks MATLAB. To conclude, the results are analyzed in § 3.3.4 comparing measurements and simulations and giving physical explanations on the observed phenomena.

3.1 Samples and setup

As discussed in the previous sections, the interest toward FTJs has grown larger and larger in the last decade and many different structures have been proposed. Although the barium titanate oxide (BTO) is not CMOS compatible and it is thus less interesting from a commercial point of view, it is widely studied due to its strong ferroelectricity down to few nanometers [58], and due to the good quality of its interfaces with metallic and semiconducting electrodes. The studied samples will be presented in the next subsection.

3.1.1 FTJ samples

All the samples investigated during this research have been fabricated at Polifab [59], the micro and nano technology center of the Politecnico di Milano, by the research group of professor Riccardo Bertacco. The chosen substrate is a commercial strontium titanate substrate ($SrTiO_3$), 100 surface, doped with Niobium (Nb) (0.5% wt.) making it electrically conductive with an electron density in the order of $\approx 10^{20}$ cm^{-3} [60]. It is an excellent substrate to epitaxially grow high-T superconductor and thin layers of many oxides [61]; at room temperature it is a centrosymmetric paraelectric material with the perovskite typical structure. Moreover, the feasibility of the growth of epitaxial layers of STO on silicon substrate through suitable buffer layers has been demonstrated [62], making STO also CMOS-compatible. The BTO epitaxial films were therefore grown by pulsed laser deposition with Nd:YAG laser with fourth-harmonic-generated wavelength of 266 nm. Other deposition conditions can be found elsewhere [63] [64]. The deposition was preceded by in-situ annealing (T= 680°C, final pressure P=40 mTorr) of the substrate template, to promote surface reconstruction, remove carbonate residuals, and provide a smooth template for optimal deposition dynamics. After the growth, BTO films, of 2 and 3 nm, were an-

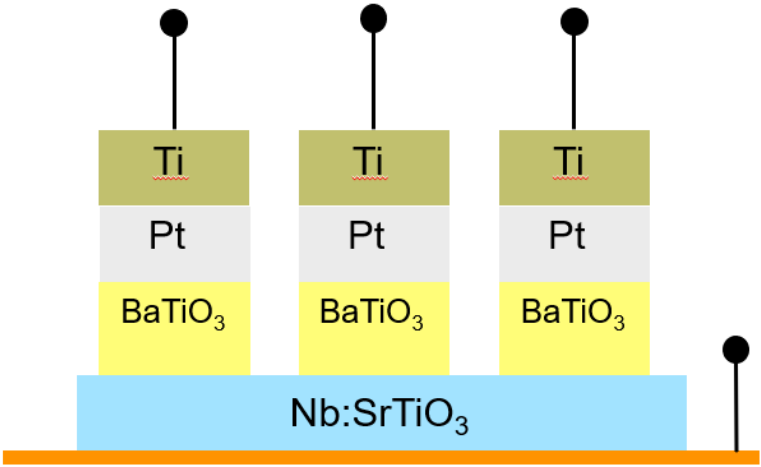


Figure 3.1: Schematic illustration of three FTJ cells on the common substrate.

nealed in situ at 600°C for 30 minutes in an oxygen pressure of 500 mbar to promote crystallization and reduce the amount of oxygen vacancies. The platinum contacts (5 nm) have been deposited ex-situ by magnetron sputtering (power = 20 W, Ar pressure = 5 mTorr). After growth, 100 nm thick Ti pillars with an area of 38×38 μm^2 were deposited by e-beam evaporation through a shadow mask on the sample. These pillars serve both as a hard mask for the subsequent etching step by Ion Beam Etchin needed to define the FTJs and as a thick metal pad for the electrical characterization. In order to be able to probe the bottom electrode, the sample was pasted on a copper sheet with silver paste. A schematic illustration of the samples is shown in Fig. 3.1, note the common substrate. To divide the various FTJ cells, Ti columns were used as rigid masks for the Ion Beam Etching process. Other details are reported elsewhere [65].

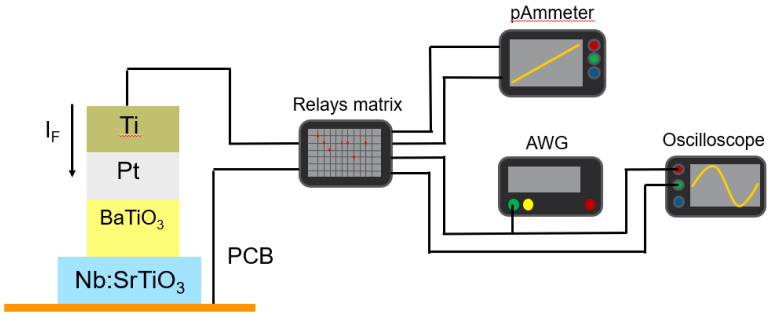
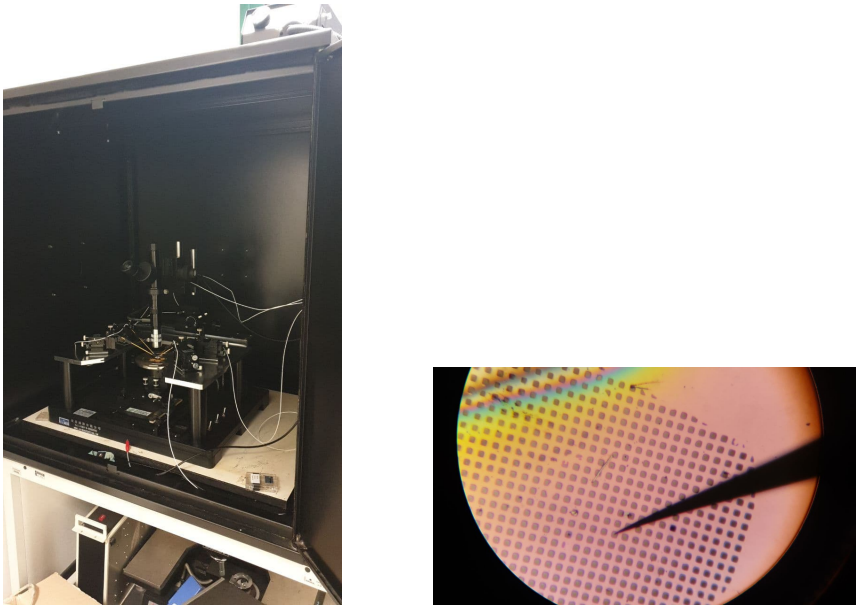


Figure 3.2: Schematic illustration of the experimental setup employed. The relays matrix allows to switch between two different setups for fast and slow measurements.

3.1.2 Experimental setup

Since the experimental investigation of these FTJ cells includes both slow and fast measurements, two different types of setup will be used. In order to switch from one to the other, without physically touching the samples and the connected cables, with the risk of altering the conditions between measurements, a in-house relays matrix is employed, controlled via an Agilent 34970A Switch Unit [66] as schematically depicted in Fig. 3.2. The two different setups are:

- **Slow or DC**: employing a Picoameter [67] that has been used for quasi-static current-voltage characteristics. It is capable of reading down to 10^{-15} A.
- **Fast or AC**: employing a TGA 12104 Arbitrary Waveform Generator [68] and a Tektronix TDS 200-Series Digital Real-Time Oscilloscope [69]. The AWG features four separate and independent channels, a sampling frequency of up to 100 MHz and a 12-bit resolution, while the oscilloscope features two channels plus one for external triggers, the channels can be set with input impedences of 1M Ω or 50 Ω. They have been used for the writing operations and to characterize time-dependent processes like



(a) (b)

Figure 3.3: (a) The probe-station structure with the micro-manipulators and the microscope. The white cables are the BNC cables for electrical connections. (b) A tungsten tip on the titanium mask of one FTJ cell, each square is the area of an FTJ 38×38 μm^2.

retention as will be described in the next sections.

The upper and the common bottom electrodes of the samples can be contacted with Tungsten tips carefully placed over the wafer surface with a couple of precision micro-manipulators and the aid of an optical microscope, as can be seen from Fig. 3.3(b). The samples and probe tips are enclosed inside a conductive structure, i.e. the probe station (Fig. 3.3(a)) which is acting as a Faraday cup for the reduction of the external electromagnetic interferences. All instrumentation was programmed using Mathworks MATLAB and connected to the computer via GPIB interfaces.

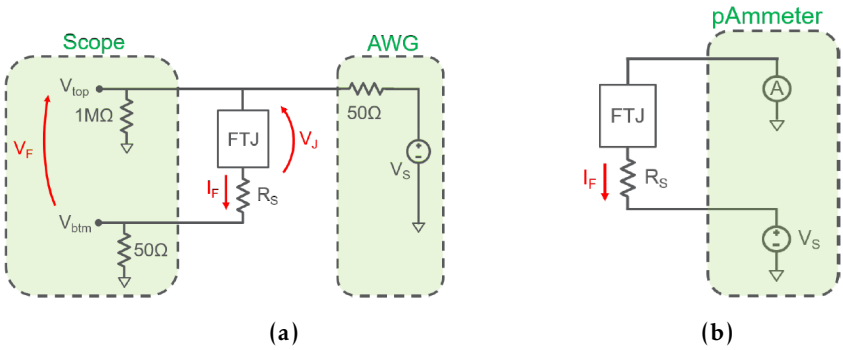


Figure 3.4: (a) Electrical model for the AC experimental setup. (b) Electrical model for the DC experimental setup.

3.2 Measurements for 1-D modeling

In the previous section the samples and the experimental setup have been presented. This section will show the results of the experimental characterization. In particular, static current voltage relation, resistive window and retention measurements will be presented with detailed explanations of the measurements procedures.

3.2.1 Static I-V

This measure employs both the AC and DC setups discussed in the previous section. The former is necessary to send the fast pulses needed to set or reset the cell under measurement (i.e. the write operation), while the latter is used for the quasi-static current measurement. In Fig. 3.4(a), the electrical model for the AC setup is shown. Pulses of positive bias will be called *set* pulses because they make the polarization vector turn toward the STO realizing the ON state or LRS. Viceversa, pulses of negative amplitude make the polarization vector point toward the metal realizing the OFF state or HRS. As reported in literature [70] [71], FTJs can be programmed into intermediate states between the most conductive state,

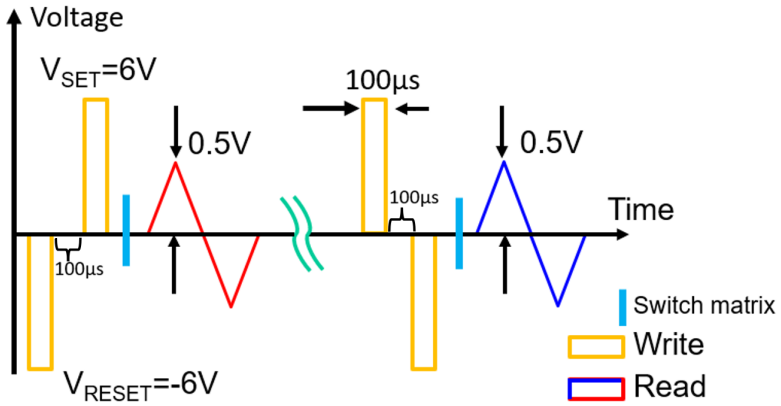


Figure 3.5: Schematic illustration of the train of pulses needed to completely reset and set the FTJ followed by I-V sensing by the pAmeter, in red for the LRS and blue for the HRS. Note that the switching matrix action is not in time scale and that the two measures (LRS and HRS) are shown together but are not necessarily performed one after the other.

called *full LRS* in the following sections, and the least conductive one, called *full HRS*. The coming I-V measures are always taken after a full set or full reset pulse. In order to do so, a preliminary analysis has been performed to discover the magnitude and duration of the pulses needed to completely set or reset the device, regardless of its initial state. The width chosen for the writing pulses is of 100 μs while the biases are $V_{reset} = -6$ V for the full reset pulse and $V_{set} = +6$ V for the full set, the pause between the two pulses is 100 μs. Using the setup shown in Fig. 3.4(a), it was possible to monitor the voltage drop between the top and bottom electrode on the oscilloscope, after which the switching matrix changes the setup to the DC one, shown in Fig. 3.4(b), for the direct measurement of the I-V curve during a 0 V$\rightarrow$ +0.5 V$\rightarrow$ −0.5 V$\rightarrow$ 0 V voltage sweep. The measurement results are shown in Fig. 3.6, where the red curve corresponds to the LRS (also in red in Fig. 3.5) while the blue curve corresponds to the HRS. The I-V curve presents 3 decades of resistive window

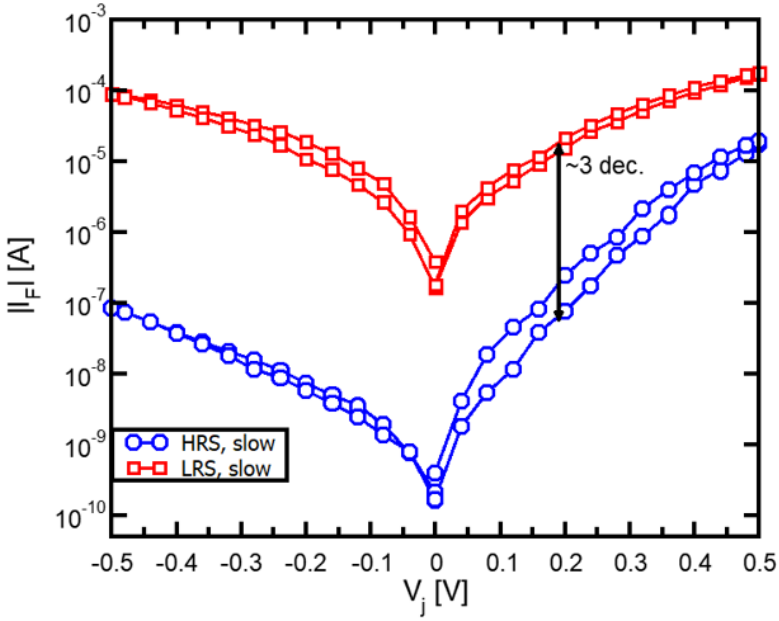


Figure 3.6: Result of the I-V measurement on the 3 nm sample. LRS in red and HRS in blue. The 3 decades of resistive window evaluated at V_j = 200 mV are highlighted.

evaluated at V_j = 200 mV. The curves are labelled as *slow* in contrast with the *fast* measurements that will be presented in the following sections. Considering all this, it is worth noting that those curves are defined as *slow* because the switching matrix takes some seconds to switch from the AC to the DC setup causing thus a relevant delay (on the FTJ switching time scale) between the programming and reading operations. As previously mentioned, note that the reading operation is non-destructive since the magnitude of the electric field reached in the ferroelectric is much lower than the coercive field needed to overwrite the programmed state.

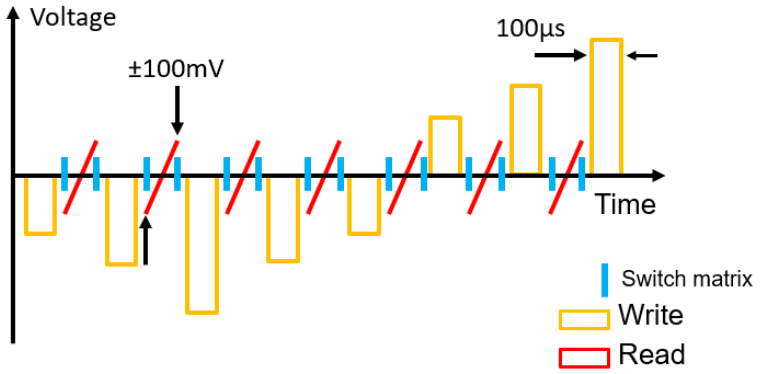


Figure 3.7: Schematic illustration of the train of pulses adopted in the hysteresis window measurement. The train of pulses starts at 0 V and the progression follow the integer numbers till $V_{min} = -6$ V, then the pulses start to increase in amplitude till $V_{max} = 6$ V. Each pulse has a width of 100 μs and between each pulse and the following one the switching matrix connect the pAmeter for the ±100 mV sweep and current sensing.

3.2.2 Resistive window hysteresis

To check the possibility of obtaining stable intermediate states as mentioned in § 3.2.1, the sample has been investigated with the so-called *hysteresis window* measurements, i.e. applying a train of pulses of increasing and decreasing amplitude, sensing the current at ±100 mV after each pulse. The measurement scheme is shown in Fig. 3.7. The result of the measurement for the 3 nm sample is depicted in Fig. 3.8. The points are obtained by averaging the current measured at ±100 mV and thus represent the low field current. As expected, the ratio between the lowest and highest current obtained with this measure is in agreement with the measure shown in Fig. 3.6, confirming that the chosen width and amplitude of the pulses is suitable for the full set and full reset operation. Moreover, it is worth noting that the pulses at V_{min} and V_{max} are sent two times but the change in the current is absent at V_{max} and almost negligible at V_{min}, giving further confirmation that the

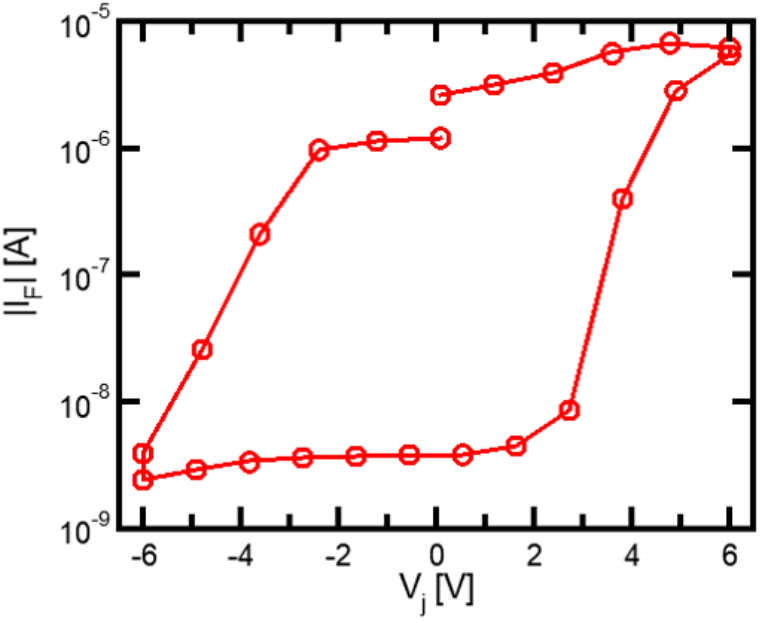


Figure 3.8: Result of the hysteresis window measurement for a 3 nm cell. As expected the ratio between the lowest and highest current obtained with this measure is in agreement with the measure shown in Fig. 3.6. It is worth noting that sending twice the pulses at V_{min} and V_{max} has negligible impact on the measured current after the pulse.

chosen intensity and duration are appropriate for achieving the so-called full set and reset. The presence of intermediate states as shown in Fig. 3.8 could have at least two different explanations [72]. The most widespread interpretation, which will be discussed in the next chapter, is that the ferroelectric is divided into domains, i.e. zones of the material with the same polarization. Adjacent domains can have opposite polarizations and thus the resulting total polarization average between domains in the LRS and domains into the HRS. The second option is that the polarization vector can change continuously between the UP and DOWN state, even though it is very unlikely that an area big as $38 \times 38\ \mu m^2$ will all be in

the same polarization state. Lastly, it is worth commenting on the slopes of the switching part of Fig. 3.8. The transition between HRS and LRS takes place at $V_j = 3$ V and it is much steeper than the transition between LRS and HRS which starts at $V_j = -3$ V. This is due to the effect of the semiconductive electrode. As shown in Fig. 2.11, the positive pulses that turn the device toward the LRS change the STO conduction band in an accumulation state and thus the major part of the applied bias falls on the ferroelectric material inducing a faster switching process (that is because in the accumulation state only a small voltage drop in the semiconductor is needed to expose a huge amount of charge). Viceversa, when the negative pulses turn the LRS into the HRS the semiconductor is in the depletion state and thus a bigger part of the applied bias drops on the depleted region of the semiconductor making the switching process slower and the curve flatter.

3.2.3 Time-resolved measurements: retention

As discussed in § 1 and § 2, the ability of a memory device to retain written data plays a key role in obtaining a reliable memory device. The measure investigating this characteristic is called *retention measurement*. In the case of FTJ, in literature, retention characteristics are usually investigated on the time scale of seconds as shown in Fig. 3.9 and Fig. 3.10. Measurement results reveal that a correct assessment of the achievable memory window in these devices cannot overlook the short-time retention loss arising from the large depolarization field [75] [76] typically present in the ferroelectric material (see § 2.3.1 and Eq. 2.7). With the aim of investigating shorter time scales we had to develop a novel, dedicated setup. The fast setup is not enough because the oscilloscope cannot read the nanometric currents of the HRS. The easiest solution is to add an amplifier in series to the device under test, but the currents generated by the programming pulses would be too high and would saturate such amplifier the mo-

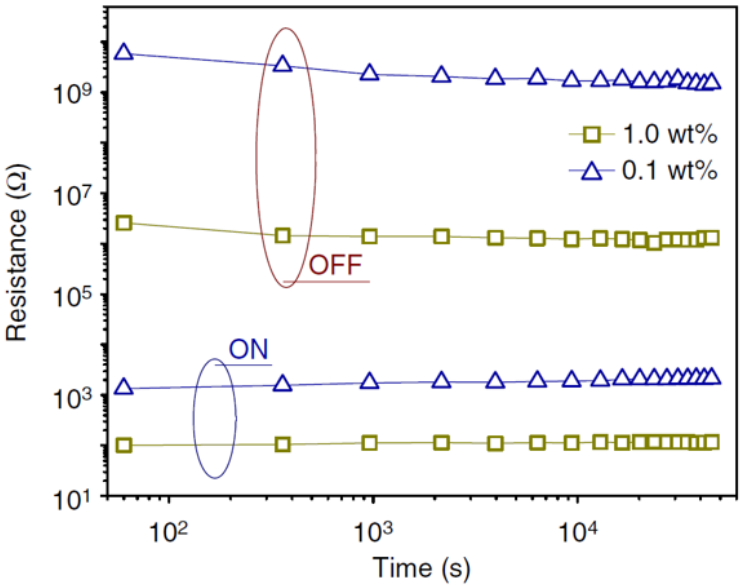


Figure 3.9: Retention measurements for a Pt/BTO/Nb:STO FTJ with two different substrate dopings, from [58].

ment immediately before the amplifier should be used to measure the read pulse currents. To overcome this issue a custom made switch board has been designed, realized and employed. Such switch (scheme in Fig. 3.12), as schematically shown in Fig. 3.11, is placed after the device under test and keeps the amplifier floating, while the programming pulses are being sent, discharging the currents into the 2^{nd} channel of the oscilloscope. As soon as the programming currents are over the switch connects the amplifier to the DUT to measure the low currents of the reading pulses.

The train of pulses employed for the retention measurements is depicted in Fig. 3.13. The switch action is illustrated in light blue, just after the set operation. For the reading operation a train of logaritmically spaced pulses of amplitude $V_j = 100$ mV is employed, sensing of I_F is performed by con-

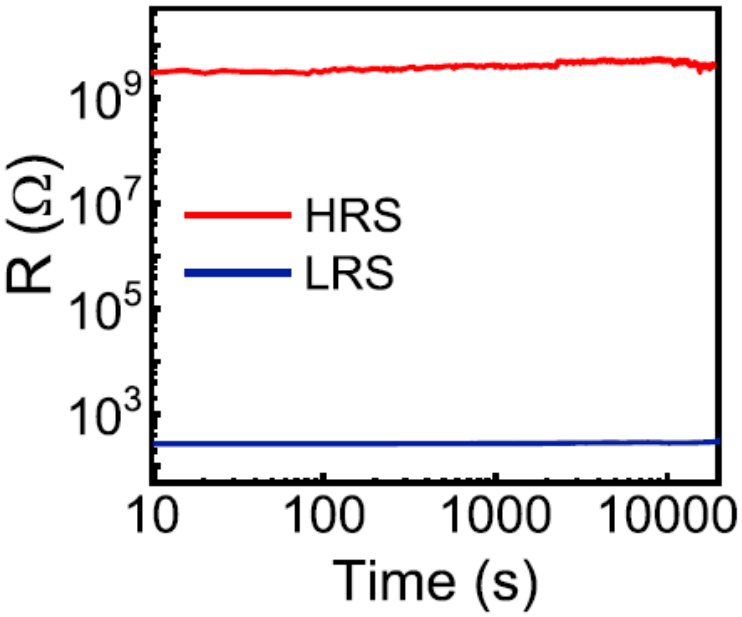


Figure 3.10: Retention measurement for a Pt/BTO/Nb:STO FTJ with 0.7 wt% substrate doping, from [56].

necting the substrate of the DUT to the TIA, whose voltage output is read by the oscilloscope and then converted back into an I_F value, the result is shown in Fig. 3.14. Note that the train of reading pulses starts 200 μs after the second programming pulse. A more detailed explanation and a physical interpretation of the phenomenon will be given in § 3.3.4. At the moment, we can just note that the LRS loses a decade of current in 100 seconds while the HRS is much more stable. Considering that the switching matrix employed to change between the fast and slow setups take some seconds to operate it is straightforward to think that such time delay may have an impact on the slow I-V characteristic shown in Fig. 3.6. To check such hypotesis the new updtated setup ha been used to measure the I-V characteristic of the LRS shortly after the programming pulse as illustrated in the next section.

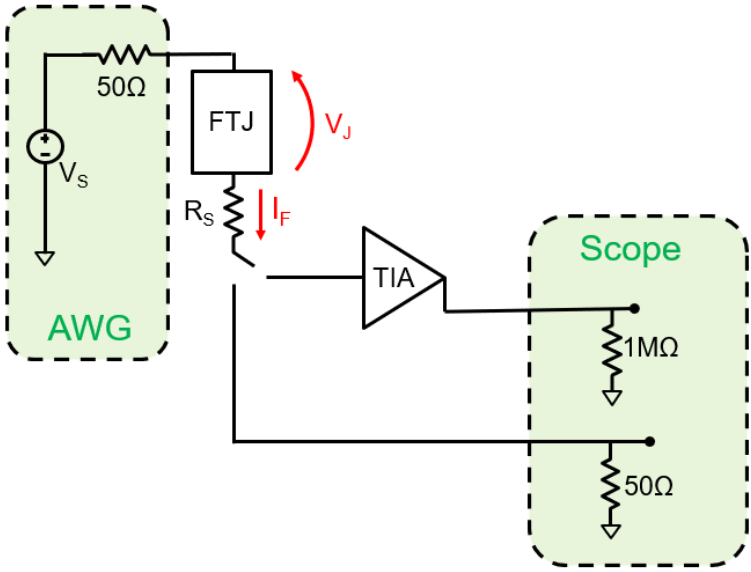


Figure 3.11: Electrical scheme of the updated fast measurement setup. The switch allows to discharge I_F into the 50 Ω of the oscilloscope to avoid the saturation of the trans-impedence-amplifier (TIA).

3.2.4 Fast I-V

The measurement scheme for the fast I-V characteristics is shown in Fig. 3.15. The reading voltage sweep 0 $V \rightarrow +0.5$ V$\rightarrow -0.5$ V$\rightarrow$ 0 V is sent 100 μs after the set voltage pulse. The current is measured by converting the bias measured on the 1^{st} channel of the scope back to the I_F by dividing for the gain of the TIA. The resulting curve is shown in green in Fig. 3.16. As expected, the I-V characteristics measured just after the programming pulse is higher than the one sensed with the picoAmeter. This is due to the retention losses acting during the relays matrix action to switch between the two setups. A further confirmation of this effect can be seen by looking at Figs. 3.17 and 3.18. From the former we can desume that the retention losses for the 2 nm sample are greater with respect

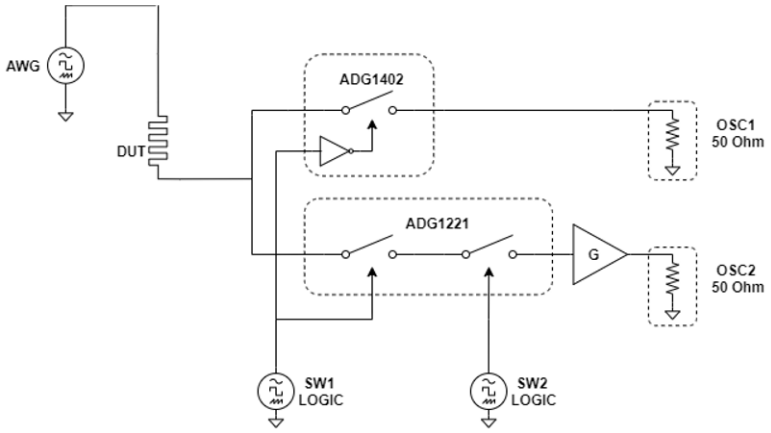


Figure 3.12: Scheme of the custom made switch board, datasheet of the devices can be found elsewhere [73] [74].

to the 3 nm sample (see Fig. 3.14) and infact the increase in the fast LRS current is bigger for the 2 nm case.

This effect is in agreement with Eq. 2.7, from which follows that a reduced thickness d brings an higher depolarization field and therefore higher losses. Moreover, this analysis also shows that, for a correct assessment of the FTJ memory window, a fast acquisition of the I-V curve of the devices is mandatory.

3.3 1-D modeling

In order to better investigate the physics behind the measurements shown in § 3.2, a custom 1-D model for electrostatics and current transport through a metal/ferroelectric/semiconductor FTJ has been devoleped in *Mathworks* MATLAB.

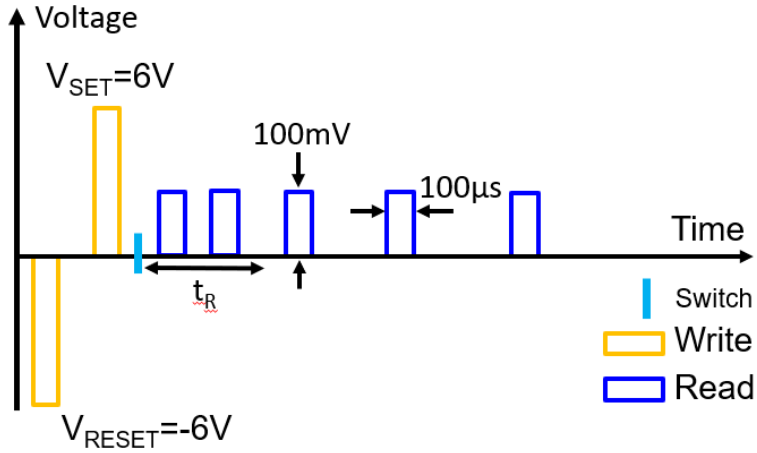


Figure 3.13: Schematic representation of the measurement scheme for the retention measurement.

3.3.1 Overview

The model solves the Poisson equation all over the device accounting for the impact of P in the ferroelectric layer by introducing two sheets of fixed charge with opposite sign at the interfaces of the material with the metal and the semiconductor, as schematically shown in Fig. 3.19. In so doing, P represents an effective 1-D polarization, averaging all the possible nonuniformities over device area. The Poisson equation is coupled with the stationary continuity equation for electrons, which includes carrier drift/diffusion and distributed tunneling in the semiconductor. Distributed tunneling is the possibility for the tunneling process of taking place not only at the interface but also through the conduction band if the region to be crossed is narrow enough [77].

3.3.2 Equations

The equations solved in the model are written explicitly in the following section. Starting with the coupled Poisson and

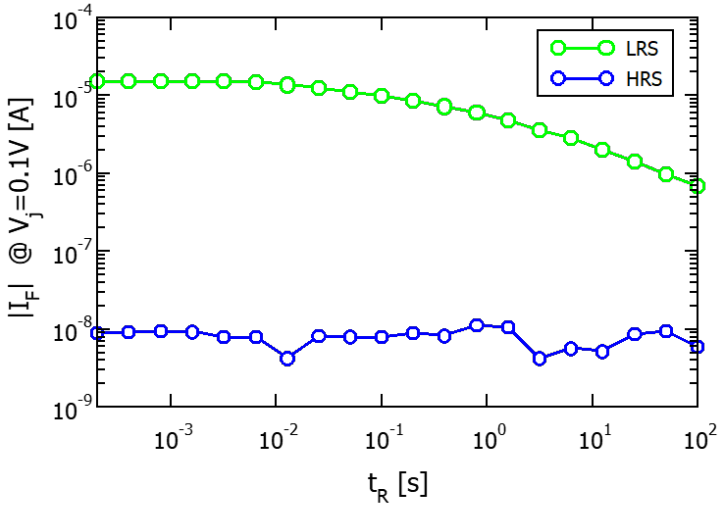


Figure 3.14: Result of the rention measurement for a 3 nm cell.

continuity equations, where the former is solved in the whole device, while the latter only in the semiconductor:

$$\begin{cases} \nabla \cdot (\epsilon \nabla \phi) = -\rho \\ \frac{\partial n}{\partial t} = \nabla \cdot \frac{J_{tot}}{q} = \nabla \cdot \frac{J_{DD}+J_{tun}}{q} = \frac{1}{q}\nabla \cdot J_{DD} - R \end{cases} \tag{3.1}$$

where q is the elementary charge, ϵ is the dielectric constant, ρ is the charge density and J_{DD} and J_{tun} are, respectively, the electron drift/diffusion current density in the conduction band and the tunneling current density in the bandgap of the semiconductor (see Fig. 3.20). ρ is defined as usual in the semiconductor, while, in order to model the polarization charges, at the oxide interfaces it takes a user defined fixed value ρ=P (defined in C/cm^2 consistent with the charge density in the semiconductor). Looking at Eq. 3.1, it can be seen that the divergence of the tunneling current has been written as a recombination term following [77] [78] [79]. Knowing

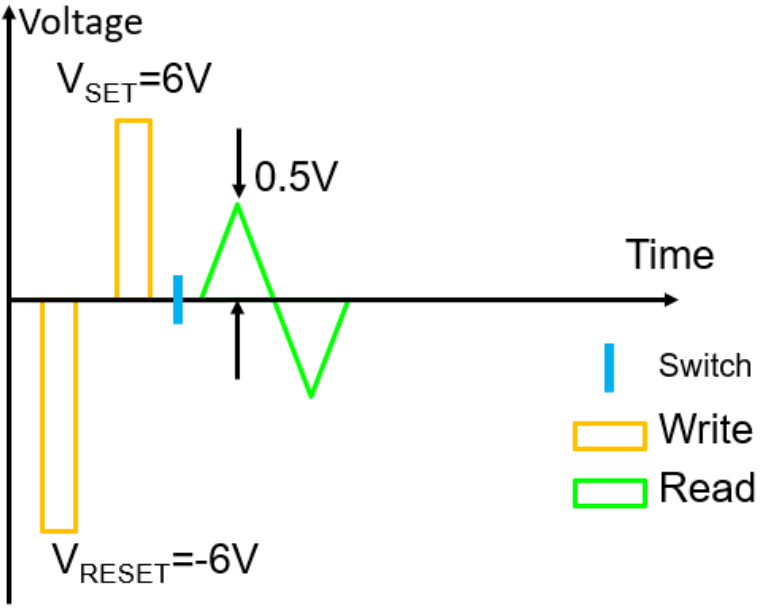


Figure 3.15: Schematic illustration of the measurement scheme for the fast I-V measurement.

that in stationary conditions the left-hand-side of the continuity equation in Eq. 3.1 is equal to 0 and that the definition of the drift and diffusion current in 1-D is:

$$J_{DD} = qn\mu_n \frac{\partial \phi_f}{\partial x} \tag{3.2}$$

the system can be rewritten in 1-D as:

$$\begin{cases} \frac{\partial}{\partial x}\left(\epsilon \frac{\partial \phi}{\partial x}\right) = -\rho(\phi_f) \\ 0 = \frac{\partial}{\partial x}\left(n\mu_n \frac{\partial \phi_f}{\partial x}\right) - R \end{cases} \tag{3.3}$$

where n is the electron density, μ_n is the electron mobility, ϕ_f is the Fermi potential in the semiconductor and R corresponds to the number of electrons (per unit time and volume) appearing/disappearing in the semiconductor conduc-

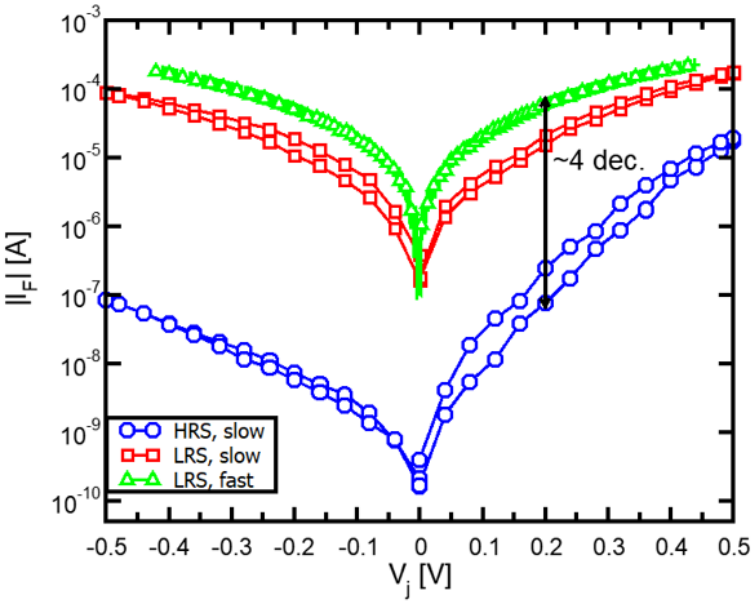


Figure 3.16: Result of the fast I-V measurement for a 3 nm cell. The fast I-V curve, i.e. the green one, is a decade higher than the slow one, achieving a total resistive window of four decades.

tion band due to tunneling at each position x and can be directly calculated as:

$$R(x) = \frac{A^* T}{k_b} F(x) \Gamma(x) \ln \left[\frac{1 + e^{[E_c(x) - E_{f,s}(x)]/kT}}{1 + e^{[E_c(x) - E_{f,m}(x)]/kT}} \right] \tag{3.4}$$

where A^* is the Richardson constant modified to account for the effective mass of the semiconductor, k_b is the Boltzmann's constant, E_C is the conduction band edge, Γ is the tunneling transparency evaluated at E_C, F is the electric field, and $E_{F,s}$ and $E_{F,m}$ are the Fermi levels in the semiconductor and in the metal, respectively. Γ was calculated under the Wentzel–Kramer–Brillouin (WKB) approximation:

$$\Gamma(E) = e^{-\frac{2}{\hbar} \int_0^{x_i} \sqrt{2 m_{tun} (E_C - E)} dx} \tag{3.5}$$

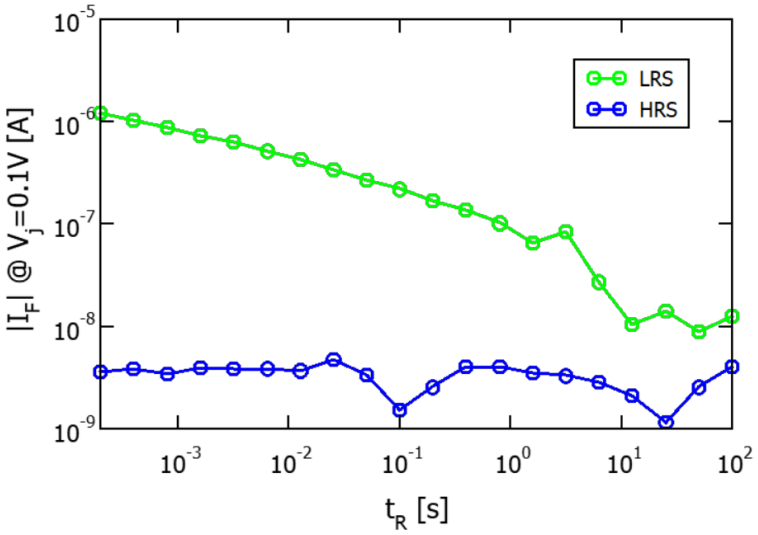


Figure 3.17: Result of the retention measurement for a 2 nm cell. The decrease of the retention current in the LRS is faster than the 3 nm case represented in Fig. 3.14, in agreement with Eq. 2.7 where it can be seen that the depolarization field increases with decreasing thickness.

whose reasonable accuracy for thin insulators was proved in [80]. The approach used is original because it applies distributed tunnelling to a junction with an oxide layer and not to a metal-semiconductor junction. Since the oxide is very thin and its energetic barrier is not too high (ΔE_C =180 meV between STO and BTO), the impact of the distributed tunneling is not negligible at all as will be shown in the fitting section.

3.3.3 Implementation

The equations shown in the previous section are solved only for the direction perpendicular to the junction, following the *finite difference* approach [81]. The simulator is written in *Mathworks* MATLAB. The meshing strategy is depicted in Fig. 3.21, all the quantities are defined on the mesh nodes, except for

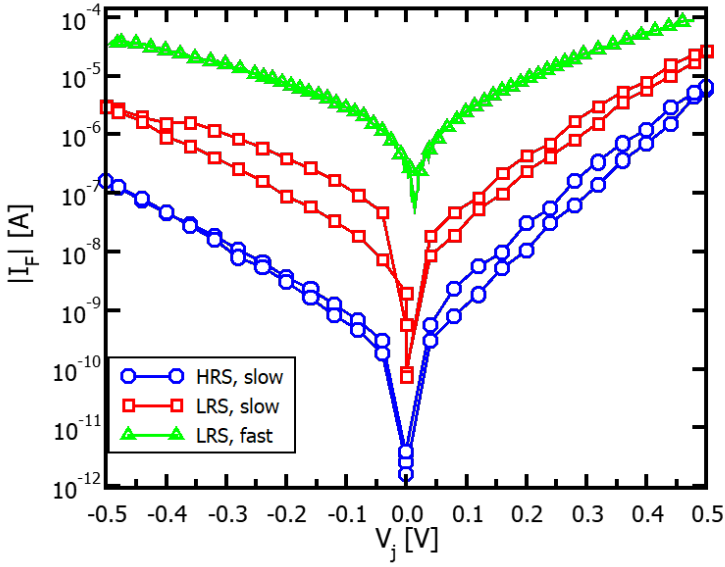


Figure 3.18: Result of the fast I-V measurement for a 2 nm cell. The increase of the LRS current with respect to the slow case is greater than the 3 nm case represented in Fig. 3.16, in agreement with Figs. 3.17 and 3.14.

the dielectric constant ϵ and the mesh step h. The Poisson equation is discretized by applying backward difference for the outer derivative:

$$\frac{\partial}{\partial x}\left(\epsilon\frac{\partial\phi}{\partial x}\right) = -\rho \rightarrow \frac{1}{h_{i-1}}\left(\epsilon_i\frac{\partial\phi_i}{\partial x} - \epsilon_{i-1}\frac{\partial\phi_{i-1}}{\partial x}\right) + \rho_i(\phi_{f,i}) = 0 \tag{3.6}$$

following by forward difference on the inner derivatives:

$$\frac{1}{h_{i-1}}\left(\epsilon_i\frac{\phi_{i+1}-\phi_i}{h_i} - \epsilon_{i-1}\frac{\phi_i-\phi_{i-1}}{h_{i-1}}\right) + \rho_i(\phi_{f,i}) = 0 \tag{3.7}$$

where ρ_i depends on ϕ_f for all the semiconductor mesh points but it is fixed in the first and last oxide mesh points $\rho_{ox,1}$=P=-$\rho_{ox,end}$. Eq. 3.7 is solved in the whole device, but in the first and last mesh point, representing respectively the top and

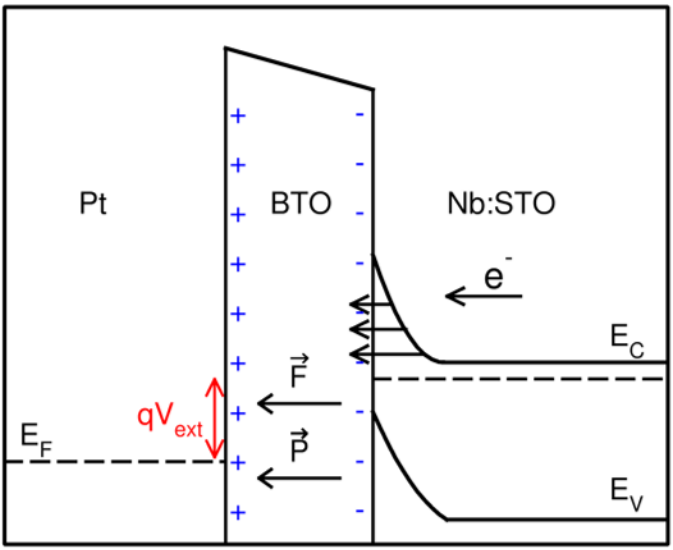


Figure 3.19: Schematic illustration of the FTJ junction in which the polarization vector is represented by two sheets of fixed charge with opposite sign at the interfaces between the ferroelectric and the metal and the ferroelectric and the semiconductor. The arrows on the side of the depleted region of the substrate represent the distributed tunneling, i.e. the possibility for the tunneling process of taking place not only at the interface but also through the conduction band, if the region to be crossed is narrow enough.

bottom electrode, where the bias is externally applied. This fact is modelled by Dirichelet boundary conditions, meaning that the potential in the first and last mesh points are not unknowns of the problem and thus do not need being computed.

The continuity equation is discretized by applying forward difference for the outer derivative:

$$\frac{\partial}{\partial x}\left(n\mu_n\frac{\partial\phi_f}{\partial x}\right)-R=0\rightarrow\frac{\mu_n}{h_i}\left(n_{i+1}\frac{\partial\phi_{f,i+1}}{\partial x}-n_i\frac{\partial\phi_{f,i}}{\partial x}\right)-R_i=0 \tag{3.8}$$

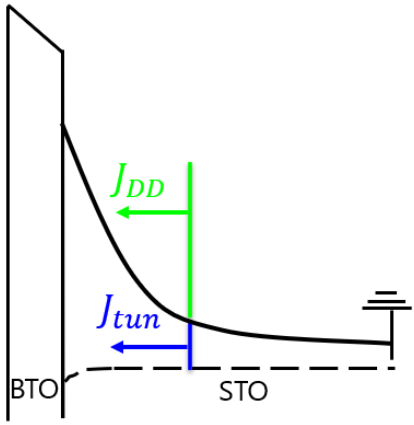


Figure 3.20: Schematic representation of J_{DD} and J_{tun}.

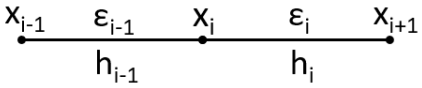


Figure 3.21: Schematic representation of the one-dimensional mesh. All the quantities are defined on the mesh nodes, except for the dielectric constant ϵ and the mesh step h.

following by backward difference on the inner derivatives:

$$\frac{\mu_n}{h_i}\left(n_{i+1}\frac{\phi_{f,i+1}-\phi_{f,i}}{h_i}-n_i\frac{\phi_{f,i}-\phi_{f,i-1}}{h_{i-1}}\right)-R_i=0 \qquad (3.9)$$

This equation is solved for all the substrate mesh points, but the last one. Such point represents the bottom electrode, where the Fermi level is imposed by the doping condition (while the bias is kept at 0 V) and thus it is modelled by a Dirichlet boundary condition and the equation is not solved for that point. A special attention must be given to the solution of the continuity equation for the first mesh node of the substrate, writing Eq. 3.9 for the first mesh point assigning $i = 1$ for

semplicity of notation:

$$\frac{\mu_n}{h_1}\left(n_2\frac{\phi_{f,2}-\phi_{f,1}}{h_1}-n_1\frac{\phi_{f,1}-\phi_{f,0}}{h_0}\right)-R_1=0 \tag{3.10}$$

but $\phi_{f,0}$ is not defined since the index 0 corresponds to a point in the ferroelectric material. This issue is solved by the *ghost mesh node* approach. At the boundary between substrate and oxide the definition of the drift and diffusion current: $J_{DD} = qn\mu_n\frac{\partial\phi_f}{\partial x} = qn_1\mu_n\frac{\phi_{f,1}-\phi_{f,0}}{h_0} = J_{DD,1}$ must be valid. Solving for $\phi_{f,0}$:

$$\phi_{f,0}=\phi_{f,1}-\frac{J_{DD,1}h_0}{qn_1\mu_n} \tag{3.11}$$

and substituting it back into Eq. 3.10 and reducing:

$$\frac{\mu_n}{h_1}\left(n_2\frac{\phi_{f,2}-\phi_{f,1}}{h_1}-n_1\frac{\phi_{f,1}}{h_0}+\frac{1}{h_0}\left(\phi_{f,1}-\frac{J_{DD,1}h_0}{qn_1\mu_n}\right)\right)-R_1= \tag{3.12}$$

$$=\left(\frac{\mu_n n_2}{h_1^2}\left(\phi_{f,2}-\phi_{f,1}\right)-\frac{J_{DD,1}}{qh_1}\right)=0 \tag{3.13}$$

that is the continuity equation for the first semiconductor node.

Calling PP the column vector, in which each element PP_i is the Poisson equation for the i_{th} element, CC the vertical vector, in which each element CC_j is the continuity equation for the j_{th} element, n_{tot} the total number of device mesh points, n_{sub} the total number of substrate mesh points and $n_u = n_{tot}-2+n_{sub}-1$ the number of unknowns, we can rewrite the system 3.3 as:

$$\begin{bmatrix}PP\\CC\end{bmatrix}=0 \tag{3.14}$$

where $\begin{bmatrix}PP\\CC\end{bmatrix}=0$ is a vector of dimsensions $n_u\times 1$. For the sake of semplicity of code writing and bug analysis the Newton method [82] has been chosen as root-finding algorithm.

Therefore, starting with guess solutions for both Poisson and continuity equations, at each iteration it can be written:

$$\begin{bmatrix} PP \\ CC \end{bmatrix} + \begin{bmatrix} \frac{\partial PP}{\partial \phi} & \frac{\partial PP}{\partial \phi_f} \\ \frac{\partial CC}{\partial \phi} & \frac{\partial CC}{\partial \phi_f} \end{bmatrix} \begin{bmatrix} \Delta PP \\ \Delta CC \end{bmatrix} = 0 \tag{3.15}$$

where $\frac{\partial PP}{\partial \phi}, \frac{\partial PP}{\partial \phi_f}, \frac{\partial CC}{\partial \phi}, \frac{\partial CC}{\partial \phi_f}$ are matrices of derivatives whose elements l, m are: $\frac{\partial PP}{\partial \phi}_{l,m} = \frac{\partial PP_l}{\partial \phi_m}$ and the column vector $\begin{bmatrix} \Delta PP \\ \Delta CC \end{bmatrix}$ is the vector of the unknown updates that will be computed at each iteration as:

$$\begin{bmatrix} \Delta PP \\ \Delta CC \end{bmatrix} = - \begin{bmatrix} \frac{\partial PP}{\partial \phi} & \frac{\partial PP}{\partial \phi_f} \\ \frac{\partial CC}{\partial \phi} & \frac{\partial CC}{\partial \phi_f} \end{bmatrix}^{-1} \begin{bmatrix} PP \\ CC \end{bmatrix} \tag{3.16}$$

The iterative loop runs till the difference between the last iteration and the previous one is under a certain tolerance for both ϕ and ϕ_f, i.e.:

$$\begin{cases} \|\phi^{k+1} - \phi^k\| < tol \\ \|\phi_f^{k+1} - \phi_f^k\| < tol \end{cases} \tag{3.17}$$

where the default tolerance is $tol = 10^{-5} k_b T/q$. The running time of the simulation is under 2 seconds for 20 biases between $V_j = -1$ V and $V_j = 1$ V with $n_{sub} = 300$. As mentioned before, even though the programming activity was very time-consuming, the custom simulator grants extreme freedom in the implementaton of physical phenomena, such as the distributed tunneling that is not always present in commercial simulation softwares.

3.3.4 Results and analysis

Figs. 3.22 and 3.23 show that the proposed model is able to successfully reproduce all the I_F–V_j curves measured on the

Table 3.1: List of parameters resulting from the calibration of the model on the measures shown in Figs. 3.16 and 3.18.

	3 nm	2 nm
t_{BTO}	3.1 nm	2.4 nm
metal work function	6.1 eV	6.1 eV
Nb:$SrTiO_3$ doping concentration	$1.6 \cdot 10^{20}$ cm^{-3}	$1.6 \cdot 10^{20}$ cm^{-3}
$BaTiO_3$ electron affinity	3.9 eV	3.9 eV
Nb:$SrTiO_3$ electron affinity	4.08 eV	4.08 eV
$BaTiO_3$ relative dielect. const	35	35
Nb:$SrTiO_3$ relative dielect. const	9	9
$BaTiO_3$ tunneling mass	$0.29m_0$	$0.55m_0$
Nb:$SrTiO_3$ tunneling mass	$15m_0$	$15m_0$
P for LRS fast	20 $\mu C/cm^2$	24.5 $\mu C/cm^2$
P for LRS slow	18 $\mu C/cm^2$	21 $\mu C/cm^2$
P for HRS	13 $\mu C/cm^2$	18.5 $\mu C/cm^2$

devices with different $BaTiO_3$ thicknesses. The list of parameters needed to match modeling and experimental results is reported in Table 3.1. The thickness of the ferroelectric layers resulting from the fit is in a nice agreement with its nominal fabrication value and the only parameters changing when moving from one device to the other are the electron tunneling mass in the $BaTiO_3$ layer and the remanent polarization P for the different device states. In this regard, it is worth noting that the values of P needed to reproduce the available data have always the same sign, corresponding to a vector pointing toward the semiconductor substrate. The origin of this may be twofold. First of all, the electric polarization of the ferroelectric layer may not be uniform over the device area, due to the partial switching of the ferroelectric domains. In a 1-D model, this makes P just an effective polarization value describing the average contribution to the tunneling current of domains with opposite polarization. Domains with upward polarization (HRS) have a much smaller weight than those with downward polarization (LRS), thus leading to an effective average polarization whose sign is al-

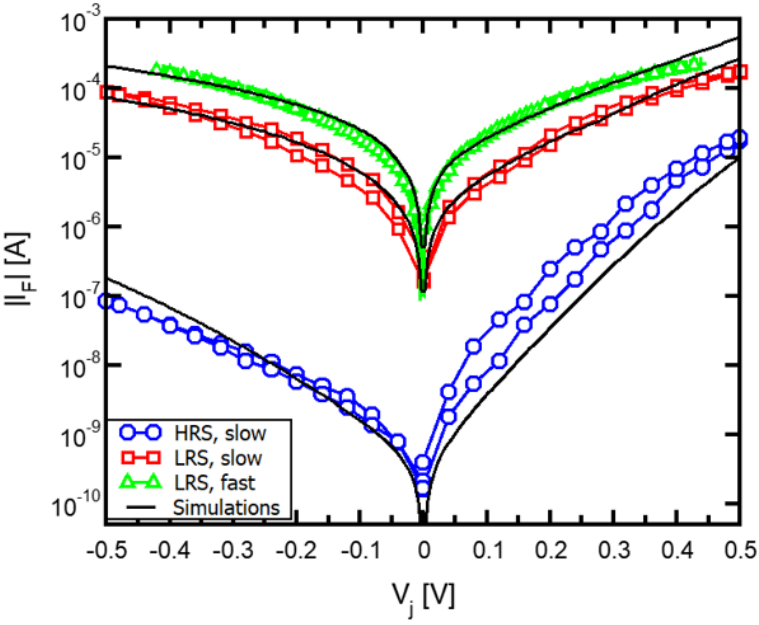


Figure 3.22: Fitting of the measure shown in Fig. 3.16 using the parameters listed in Table 3.1. Note that the limitation of V_j to 0.5 V during the voltage sweeps used to monitor I_F makes the hysteresis between the forward and reverse curves negligible.

ways determined by domains with inward P. Besides, the values of P may reveal the presence of some spurious positive charge trapped close to the Nb:SrTiO3/BaTiO3 interface and not switched by the set/reset pulses. These points will be the subject of future further investigation. The change of the electron tunneling mass for the BaTiO3 layer in the devices with different t_{BTO} values, on the other hand, may be a consequence of the change of the microscopic properties of the material with its thickness, similar to what typically reported for nanoscale SiO2 layers [83] [84]. Fig. 3.24 shows the calculated conduction-band profile of the device for the 3 nm sample in the case of $V_j = 0$ V, under LRS-fast [Fig. 3.24(a)] and HRS conditions [Fig. 3.24(b)]. Results reveal, first of all,

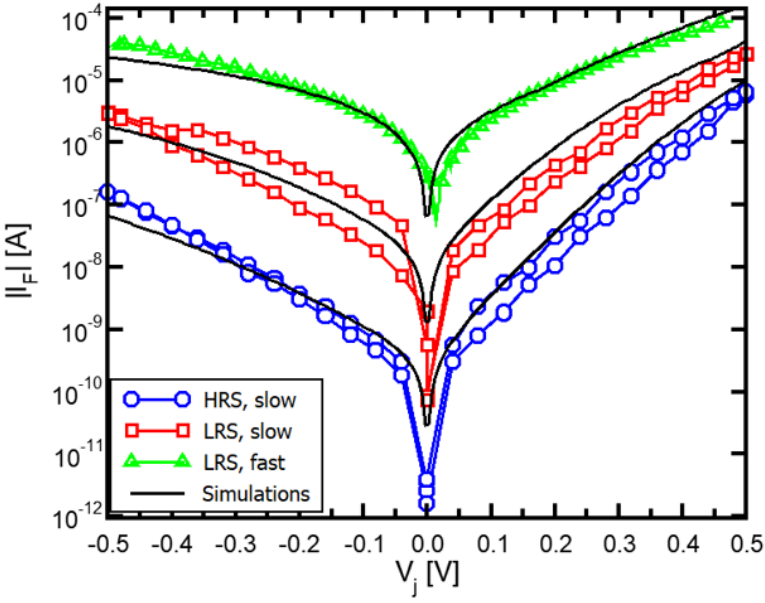


Figure 3.23: Fitting of the measure shown in Fig. 3.18 using the parameters listed in Table 3.1. Note that the limitation of V_j to 0.5 V during the voltage sweeps used to monitor I_F makes the hysteresis between the forward and reverse curves negligible.

that the reduction of P when moving from the LRS to the HRS creates a depletion layer, and hence strongly reduces the electron concentration, at the substrate surface. This is in agreement with and validates through careful quantitative modeling, the physical picture proposed in [55] and [58] that metal/ferroelectric/semiconductor FTJs may exploit the modulation of the band bending and the electron concentration at the substrate surface to achieve memory windows wider than those of metal/ferroelectric/metal FTJs. In addition to that, the band diagrams of Fig. 3.24(a) and (b) highlight that the change of P moving from the LRS to the HRS does not allow to change the direction of the electric field and reverse the band bending in the $BaTiO_3$ layer. This is due to a

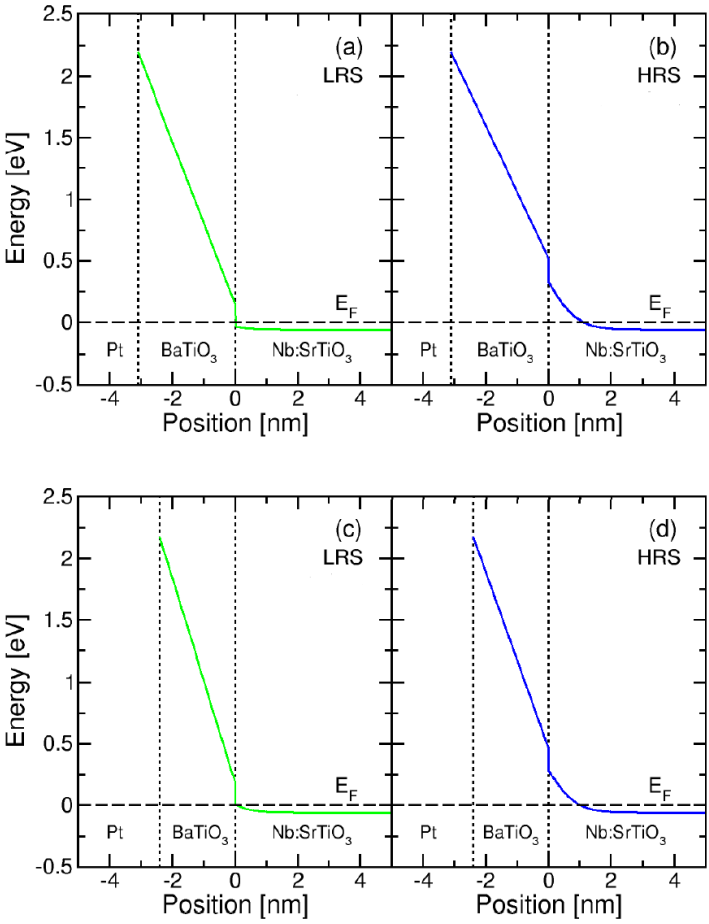


Figure 3.24: Computed conduction band profile at $V_j = 0$ for the device from (a) and (b) 3 nm sample and (c) and (d) 2 nm sample, in the (a)-(c) LRS-fast and (b)-(d) HRS.

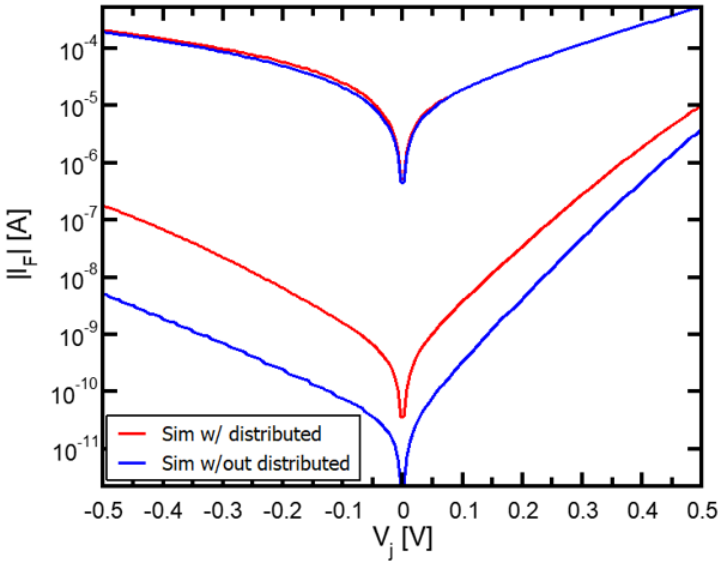


Figure 3.25: Calculated I_F-V_j curves of the 3 nm FTJ when considering and when neglecting the distributed tunneling through the semiconductor.

relatively large built-in electric field in this layer coming from the mismatch between the work function of the electrodes, which is not overwhelmed by the electric field contribution arising from the polarization charge. The direction of the band bending in the ferroelectric layer, besides, is expected to make P rather stable in the HRS. This is in agreement with the stability of I_F in the HRS curves of Figs. 3.14 and 3.17. On the other hand, when the device is in the LRS, the direction of the band bending in the $BaTiO_3$ layer is unfavorable to the stability of P. Considering that the polarization charge in that state increases the electric field and the band bending in the ferroelectric layer, as clearly evident when comparing Fig. 3.24(a) and (b), this should result in a strong tendency of P to reduce its value over time. This is in agreement with the relevant drift of I_F in the LRS curve in Figs. 3.14

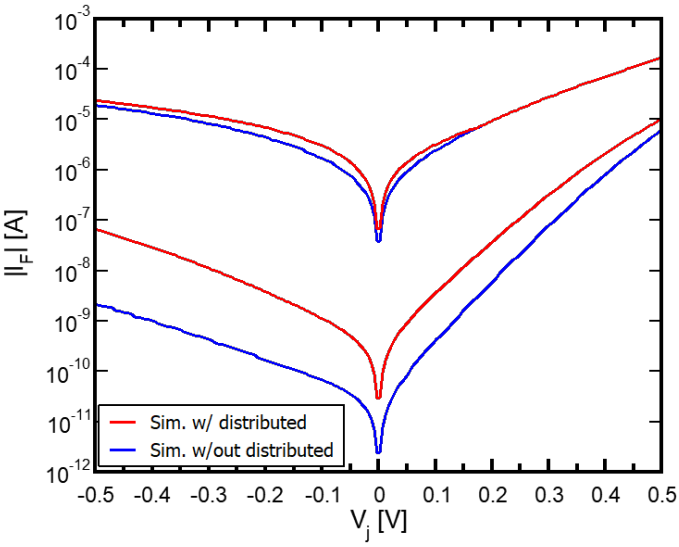


Figure 3.26: Calculated I_F-V_j curves of the 2 nm FTJ when considering and when neglecting the distributed tunneling through the semiconductor.

and 3.17. Within this picture, the even faster decrease of I_F in the LRS exhibited by the device from the 2 nm sample in Fig. 3.17 can be explained by considering that the band diagrams for this device predict a stronger electric field and band bending in the $BaTiO_3$ layer with respect to the case of the 3 nm sample, as appearing when comparing Fig. 3.24(a) and (c). To conclude this discussion, it is worth mentioning that the previous results suggest that the improvement of the time stability of the HRS and LRS of an FTJ requires a careful design of the electric field that is present in the ferroelectric layer during data retention. In this regard, engineering the built-in field originating from the work-function difference of the electrodes is mandatory, as recently pointed out in [85]. Adopting a ferroelectric material with a high coercive electric field may also help. From this standpoint,

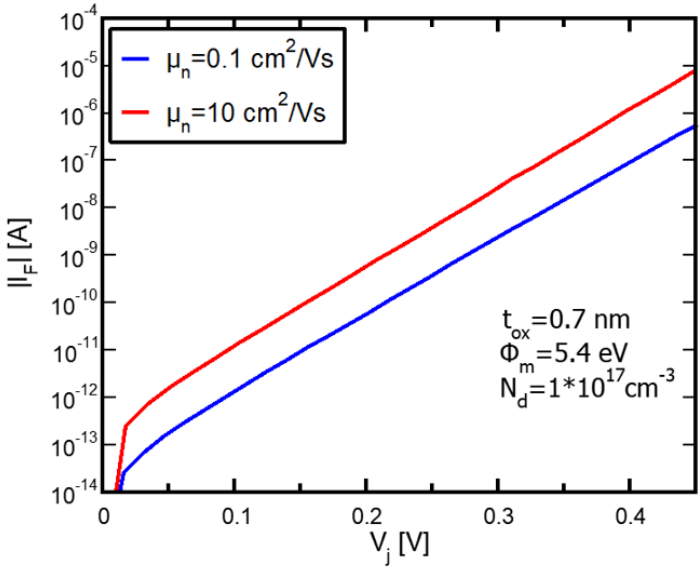


Figure 3.27: Calculated I_F–V_j curve of an FTJ with t_{BTO}=0.7 nm, substrate doping concentration equal to $5 \cdot 10^{18}$cm^{-3}, P=13.5 μC/cm^2, and metal work function equal to 5.4 eV, in the case of electron mobility equal to 10 cm^2/V·s and 0.1 cm^2/V·s. All the other parameters are the same as reported in Table 3.1.

ferroelectric-HfO_2 represents a promising solution for future high-performance and highly reliable FTJs due to its higher coercive electric field and lower dielectric constant with respect to perovskite ferroelectrics [86]. In addition, the high quality of HfO_2 may also help improving data retention at long times when charge trapping in the material is expected to play a relevant role [85] [87]. The good agreement between modeling and experimental results in Figs. 3.22 and 3.23 allows to exploit the proposed model for current transport through metal/ferroelectric/semiconductor FTJs to investigate specific points related to the physical processes involved in the electron flow. The first of these points is the impact of distributed tunneling through the semiconductor

on the total I_F. In this regard, it is worth noting that the creation of a depletion layer, and its consequent band bending, at the semiconductor surface in the HRS allows to strongly reduce the tunneling electron flow between the metal and the semiconductor interfaces. This is what allows FTJs with a semiconductor electrode to achieve I_F windows that are much wider than those achievable by FTJs with both electrodes made of a metal. However, the adoption of a high doping concentration in the semiconductor to reduce its parasitic resistance limits the width of the depletion layer so much that distributed tunneling through it introduces a relevant contribution to the total I_F flowing between the FTJ electrodes. This, in turn, constrains the achievable memory window of the device. Such reduction of memory window is clearly observable in Figs. 3.26 and 3.25, where the I_F–V_j curves for the LRS-fast and HRS conditions of the device from the 3 and 2 nm samples, calculated with and without distributed tunneling in the semiconductor are reported (in the latter case G_{tun} was set to 0 for all the positions except for the interface). Results reveal that when the device is in the HRS, neglecting distributed tunneling in the semiconductor determines a reduction of I_F between one and two orders of magnitude depending on V_j and, in turn, the width of the depletion layer. A weaker change of I_F appears, instead, when the device is in the LRS, since in that case, the semiconductor surface is not significantly depleted of carriers (see the band diagrams of Fig. 3.24) and the tunneling process is mainly an interfacial phenomenon. A second point related to current transport that can be addressed by the proposed model for FTJ operation is the constraint to I_F coming from the drift/diffusion of charge carriers along the semiconductor. In this regard, it is worth mentioning that the drift/diffusion process may constrain the electron flow in the device mainly depending on the value of the following parameters: 1) electron mobility in the semiconductor; 2) doping concentration of the semi-

conductor; and 3) tunneling transparency of the ferroelectric layer (due to current continuity). In the case of the samples investigated here, a Nb:$SrTiO_3$ substrate with a high doping concentration close to 10^{20} cm^{-3} and an electron mobility close to 10 $cm^2/V{\cdot}s$ was used. This fact, in the presence of a $BaTiO_3$ layer with thickness of a few nanometers (and with the parameters reported in Table 3.1), made drift/diffusion along the substrate a negligible constraint to the current flow through the devices. This means that the I_F–V_j curves calculated considering I_F as a pure tunneling current between the edge of the quasi-neutral region of the substrate and the metal are identical to those reported in Figs. 3.23 and 3.22 (no constraint from in-band transport in the semiconductor). At any rate, this may not be true if a different device design or different materials were adopted. For instance, the reduction of the doping concentration of the semiconductor with respect to our samples was shown to enhance the memory window of the devices due to stronger depletion effects in the HRS [58]. Besides, very thin dielectric layers may be used to give rise to crested tunneling barriers allowing to optimize device performance [51]. These solutions may create conditions under which the role played by drift/diffusion of carriers in the semiconductor cannot be neglected, especially if materials with very low electron mobilities are used for the substrate [88]. This is shown in Fig. 3.27, where the simulation results for I_F with t_{BTO}=0.7 nm, substrate doping concentration equal to $5 \cdot 10^{18}$ cm^{-3}, P=13.5 $\mu C/cm^2$, and metal work function equal to 5.4 eV are shown, in the case of electron mobility equal to 10 $cm^2/V{\cdot}s$ and 0.1 $cm^2/V{\cdot}s$. The decrease in I_F following the mobility reduction reveals a non-negligible constraint from carrier drift/diffusion in the semiconductor electrode. This shows that the proposed model offers the most comprehensive description of current transport in a metal/ferroelectric/semiconductor FTJ presented so far, including not only the effects of the modulation of the

tunneling transparency of the ferroelectric layer (as done in many previous works([89] [75] [90] and [91]) but also of distributed tunneling (similar to what done in [92] and [93]) and of carrier drift/diffusion in the semiconductor electrode (never accounted for before). To conclude, it is worth mentioning that the model presented and validated here for a $Pt/BaTiO_3/Nb:SrTiO_3$ metal/ferroelectric/semiconductor FTJ can readily be applied to device structures more compatible with the CMOS process flow. For instance, moving to the widely investigated metal/HfO_2/SiO_2/silicon structure requires just to add the interfacial SiO_2 dielectric in-between the semiconductor electrode and the ferroelectric layer in the simulation domain and to modify the properties of the materials reported in Table 3.1. Nothing changes, instead, in the simulation flow used to come to the self-consistent solution of the Poisson equation and of the continuity equation for electrons and allowing to come to I_F.

3.3.5 Conclusions

In this chapter, the experimental characterization and 1-D modeling of current transport in FTJ devices have been presented. First, the experimental setup and the description of the FTJ samples are exposed, followed by the presentation of the measurements schemes and the relative results. In particular device behavior at short times from the set/reset operations using a fast measurement setup is presented for the first time. Secondly, the developed model, the simulation algorithm and the detailed equations, are shown and the distributed tunneling feature is explained. In the next sections, the measures are fitted with the developed simulator, showing that the distributed tunneling is an essential feature for this kind of devices. The huge polarization loss at short time in the LRS is explained with the aid of the simulator, showing the electronic band profiles of the LRS and HRS and showing how the electric field is almost always negative inside the

ferroelectric layer during retention. Moreover, the impact of very low mobility materials is investigated through simulations, showing the potential of the proposed model.

The developed model has the advantage of being farly simple and thus the simulator code could be easily adapted to other kind of structure (e.g. crested barrier FTJs), but, as recently pointed out, the actual physical picture may be much more complex for such structures [94] [95]. Another important fact to point out is that this thesis focuses on direct and distributed tunneling as transport mechanisms in order to highlight the possible impact of the latter, neglecting other contributions that may be important, especially for thicker ferroelectric layers, such as trap or defect assisted tunneling [96] [97] or Poole-Frenkel transport [98] [99]. Nevertheless, this simulator may represent an useful tool both for device physics investigation and FTJ parameters optimization.

Chapter 4

Ferroelectric switching modeling

This chapter presents modeling activities on FTJ ferroelectric switching. In § 4.1 a brief introduction to the phenomenon of ferroelectric switching with a historical overview of the models used is found. Throughout § 4.2 the additions to Landau's theory needed for our model are presented: the interactions between adjacent domains and the time evolution description in the diffusive limit, i.e. the Landau-Khalatnikov equation. In § 4.3, stemming from a recent work on ferroelectric capacitors, the developed model for the switching simulation is described in detail. In the end the early results are presented in § 4.3.2 comparing measurements and simulations and giving possible interpretations on the results, along with possible future development.

4.1 Introduction

The switching in a ferroelectric material is the process by which the polarization vector P changes direction. The theory of switching has its roots in the study of phase transitions in the 1930s [100] [101] [43]. One of the fundamental ideas of the switching theory is the concept of the *ferroelectric domain*. A ferroelectric domain is an area of oriented spontaneous polarization, i.e. an area in which all the dipols contributing to the macroscopical polarization are parallel to one another. There are many reasons for the existence of domains, including nonuniform strain and microscopic defects, but even in an ideal crystal, domains are to be expected for energetic reasons. In a bulk material we can have many domains with different polarization directions, the macroscopic polarization is the weighted sum of all the domains polarizations (see Fig. 4.1). In principle, P may assume any direction in a bulk material, but this may not be the case for an epitaxial thin film where it has been shown that the out of plan direction is enhanced [102] [103]. Many mathematical models to describe ferroelectric switching have been proposed: the traditional one, often called the Kolmogorov-Avrami-Ishibashi (KAI) model, is based on the classical statistical theory of nucleation and unrestricted domain growth, whereas the time dependent change in polarization is given as:

$$\Delta P(t) = 2P_s[1 - e^{-(t/t_0)^n}] \tag{4.1}$$

where n and t_0 are respectively the effective dimension and characteristic switching time for the domain growth. This model has been improved by adding a Lorentian distribution of the characteristic switching time [104], but it is not always accurate [105]. Other models have been proposed, such as the nucleation-limited-switching (NLS) [106], in which, in contrast to Kolmogorov-Avrami-Ishibashi approach, it is assumed that the film consists of many areas, which have independent switching dynamics. The switching in an area is

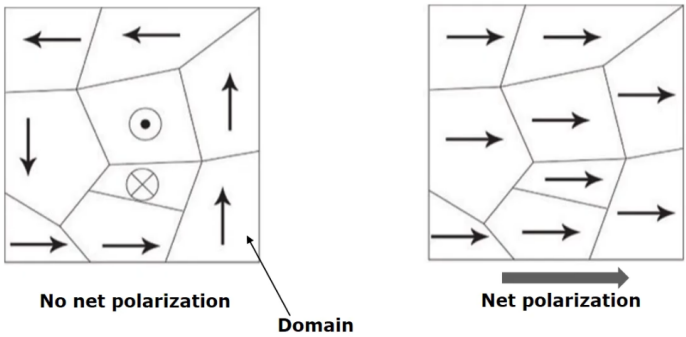


Figure 4.1: Schematic illustration of domains in a ferroelectric material. On the left the domains polarization do not give rise to any net polarization, viceversa, on the right, they are parallel to one another and give rise to a net polarization pointing to the right.

considered to be triggered by an act of the reverse domain nucleation and the switching kinetics is described in terms of the distribution function of the nucleation probabilities in these areas. These models describe the switching process within a statistical framework of areas and switching time; throughout this chapter another approach, based on the Landau equation, will be presented.

4.2 Landau-Devonshire-Ginzburg

Landau theory is based only on simmetry considerations and can provide a reliable description of a system's equilibrium near a phase transition. Since the theory assumes spatial averaging of all local fluctuactions, it is particularly suitable for ferroelectrics due to their long range interactions. The transition is characterized in terms of an order parameter, which is the polarization P for a ferroelectric system. The free energy F_P in the vicinity of the transition is expanded as a power series of the order parameter where only simmetry-compatible terms are retained. To find the state of the system the free en-

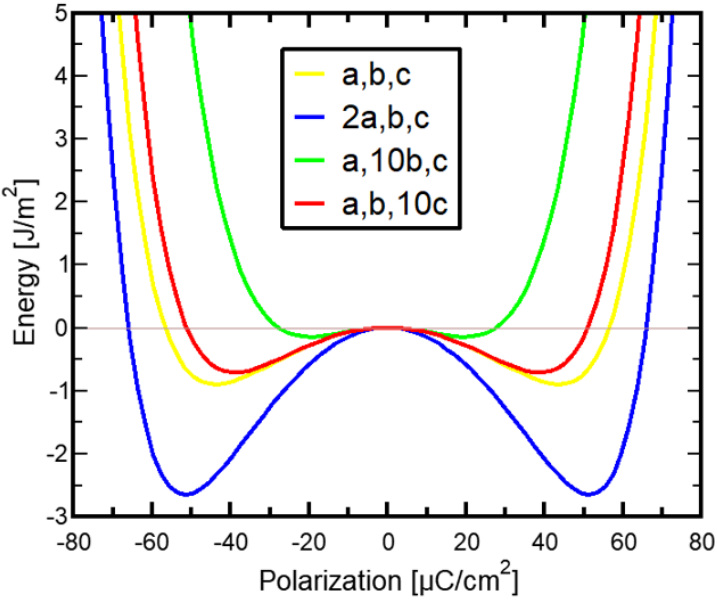


Figure 4.2: Plot of Eq. 4.2 for various Landau parameters and $E = 0$. The two wells are the stable polarization configurations.

ergy is then minimized leading to the stable spontaneus polarization. The expression for the free energy can be written as:

$$F_P = \frac{1}{2}a(T - T_C)P^2 + \frac{1}{4}bP^4 + \frac{1}{6}cP^6 + \frac{1}{8}dP^8 - EP \qquad (4.2)$$

where a, b, c, d are the Landau coefficients that can either be measured or computed with first principle calculations, T_C is the Curie temperature and E is the electric field inside the ferroelectric. A plot of the equation with E=0 is shown in Fig. 4.2, where the two wells represent the opposite polarization values of the stable configuration. Once an electric field is applied one of the two configurations is preferred with respect to the other one, i.e. its energy is lower as shown in Fig. 4.3. If the electric field is lower than the coercive field,

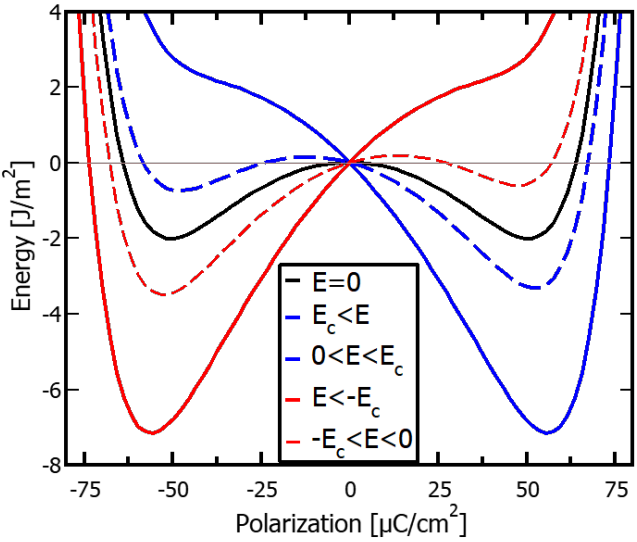


Figure 4.3: Effect of the electric field on the plot shown in Fig. 4.2.

one direction is favoured but if the system is already in the other state it will keep its state (e.g. the positive local minimum of the red dashed curve in Fig. 4.3). Viceversa, if $E < -E_c$ or $E > E_c$ the system will be forced into the only minimum left (e.g. the positive well of the solid blue curve in Fig. 4.3).

In order to take into account the spatial variations of P, and thus the domain wall contributions, the so-called Ginzburg term is added:

$$F_P = \frac{1}{2}a(T - T_C)P^2 + \frac{1}{4}bP^4 + \frac{1}{6}cP^6 + \frac{1}{8}dP^8 - EP + k|\nabla P|^2 \quad (4.3)$$

where k is the domain coupling constant.

4.2.1 Landau-Khalatnikov equation

In the previous section the Landau-Devonshire-Ginzburg has been presented. With the aim of describing the temporal evolution of the system we must introduce the Landau-Khalatnikov

equation [107]:

$$-\frac{\partial F_P}{\partial P} = \rho \frac{\partial P}{\partial t} \tag{4.4}$$

where ρ is a damping parameter of unit [Ωm], which is a measure of the loss in the material. It is worth noting that Eq. 4.4 cannot describe the intrinsic dynamics during ferroelectric switching but is only a diffusive limit approximation. In this framework, the dynamics of the system is described by the domain wall velocity, which depends on the parameter ρ. To overcome the diffusive limit molecular dynamics simulations would have to be carried out.

The coupling of Eqs. 4.3 and 4.4 takes to the same formulation of the time-dependent Ginzburg-Landau equation which is known from the theory of superconductivity [108], linking the time evolution of the polarization to the derivative of the free energy with respect to the polarization itself.

4.3 2D modeling

In the previous sections a very brief introduction of the Landau description of the ferroelectric switching has been given. By coupling Eqs. 4.3 and 4.4 the time evolution of the polarization vector of a ferroelectric domain is described along with the impact of the adjacent domains. In order to describe the switching of the ferroelectric layer in an FTJ device the approach used here follows that of Hoffman's work [109]. In that paper, about ferroelectric capacitors, the inhomogeinity of P on the device surface was modelled by a scattering of the Landau parameters over the area and the depolarization effects were neglected due to the thickness of the considered sample, along with the leakage (tunneling) current. These assumptions were made to simplify the analytical description of the device electrostatics and to be able to compute every quantity through equations simple enough. Such assumptions are clearly inapplicable to the FTJ sample, since the de-

polarization effects are stronger when the ferroelectric layer is thinner and since the tunneling current is the main feature of the device.

The assumption of perpendicular polarization will be maintained, while the one-dimensional model presented in § 3.3, being selfconsistent, can overcome the limits of the Hoffman approach, being able to compute the field inside the ferroelectric material and the current flowing through the device in any condition, given that the analyzed state is a steady one and not a transient. It should be noted that the resulting approach can be called *quasi2-D*, since, as will be presented in the next subsection, the sole bidimensional equation is Eq. 4.3 while current and electrostatics are computed only in the perpendicular direction.

The aim is running a simulation of the measure shown in § 3.2.2. In that measure, a train of pulses of 100 μs with increasing amplitude were sent in order to program the device and the mean current at ±100 mV was measured after each pulse followed by a brief pause needed to commute the switching matrix (see the measurement scheme in Fig. 3.7). The current will be therefore simulated only after the switching process has been completed with the application of the high amplitude writing pulses. The area will be discretized with a two-dimensional mesh and also the time will be discretized. In the following subsections the model implementation will be described in detail.

4.3.1 2D implementation

As briefly discussed in the previous subsection, the starting point is the discretization of the area into a 2-D mesh where each element is a square with a side of 380 nm, resulting in a 100×100 grid. Then, to each element of this grid is assigned a set of Landau parameters. Each parameter is randomly chosen from a Gaussian with fixed mean and standard deviation. It is worth underlining that these elements are not

ferroelectric domains but an artifical discretization to model the non-uniformity of the ferroelectric layer. The next step is to combine and discretize Eqs. 4.3 and 4.4:

$$\begin{cases} F_P = \frac{1}{2}a(T-T_C)P^2 + \frac{1}{4}bP^4 + \frac{1}{6}cP^6 + \frac{1}{8}dP^8 - EP + k|\nabla P|^2 \\ -\frac{\partial F_P}{\partial P} = \rho\frac{\partial P}{\partial t} \end{cases} \tag{4.5}$$

obtaining:

$$\rho\frac{\partial P}{\partial t} = -(a(T-T_C)P + bP^3 + cP^5 + dP^7 - E - 2k\nabla^2 P) \tag{4.6}$$

Following the forward Euler scheme for the time discretization we know:

$$\rho\frac{\partial P}{\partial t} \approx \rho\frac{P_{t+1} - P_t}{\Delta t} \tag{4.7}$$

where P_{t+1} and P_t are the values of P at the time instants $t+1$ and t respectively. We can write the polarization at the next time step, knowing the one at the previous time step:

$$P_{t+1} = P_t - \frac{\Delta t}{\rho}(a(T-T_C)P + bP^3 + cP^5 + dP^7 - E - 2k\nabla^2 P) \tag{4.8}$$

This equation must be solved for each square of the mesh grid, for each time step. By defining the matrix P on the mesh grid with elements $P_{i,j}$ we can write explicitly the discretization of the second derivative on the area for each element $P_{i,j}$:

$$\nabla^2 P_{i,j} = \frac{P_{i+1,j} + P_{i,j+1} + P_{i-1,j} + P_{i,j-1} - 4P_{i,j}}{\Delta x \cdot \Delta y} \tag{4.9}$$

thus reducing the gradient interaction only to the nearest neighbours of each square of the mesh. In the boundary points the contributes from the point exceeding the mesh are neglected, e.g.:

$$\nabla^2 P_{100,100} = \frac{P_{99,100} + P_{100,99} - 4P_{100,100}}{\Delta x \cdot \Delta y} \tag{4.10}$$

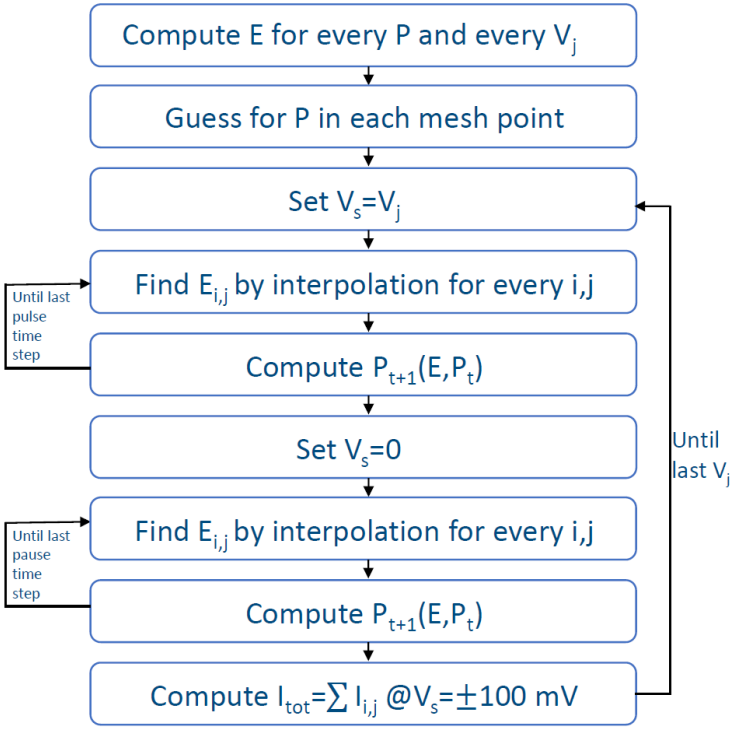


Figure 4.4: Simulation scheme.

since $P_{101,100}$ and $P_{100,101}$ do not exist.

With Eq 4.8 the polarization at the next time step for each mesh point can be computed knowing the polarization in that point and its four nearest neighbours at the previous time step. The other needed parameter is the electric field in the considered mesh point, that is a function of P and it is computed through the 1-D simulator presented in § 3.3. The computation of E and P_{t+1} must be performed $100 \cdot 100 = 10^4$ times for each time step. If the time discretization Δt is small with respect to the total time of the switching simulation the computation can become very time consuming, not to mention the fact that the above described operations are needed for the switching simulation for a single writing pulse, while in

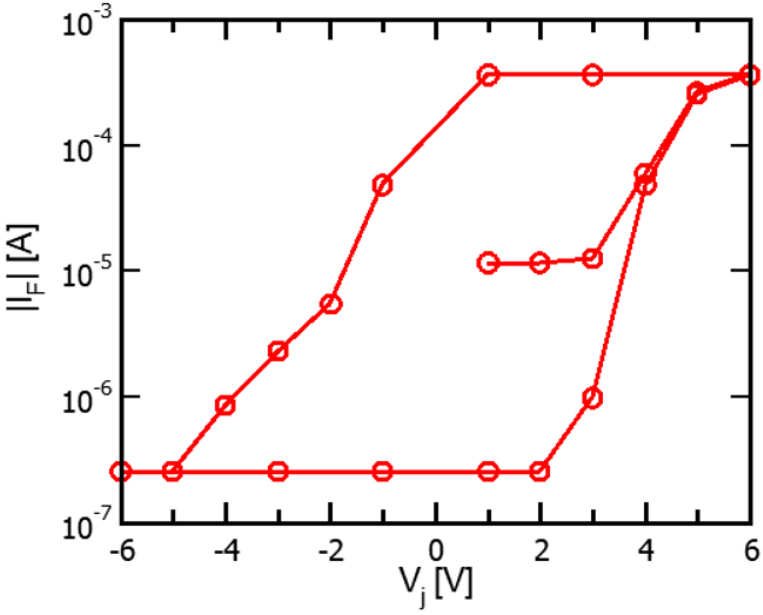


Figure 4.5: Simulated resistive window using parameters listed in Table 4.1.

Fig. 3.8 there are 24 points corresponding to as many writing pulses.

It is worth noting that to reduce the simulation time the electric field does not depend on the Landau parameters set that are different for each mesh point. The electric field E is a function of all the 1-D parameters listed in Table 3.1, that are equal for all the mesh points, the polarization P and the applied bias V_j. Since the applied biases and the amplitude of the writing pulses are fixed, and since the values P can take are limited within a certain range (see Fig. 4.2 no value above 80 μC/cm^2 is permitted) the computation of the electric field E for each mesh point and each time step can be avoided. Such computation can be done in advance for each V_j and each P and the resulting electric field can be stored in a matrix. The matrix will have a row for each V_j and a column

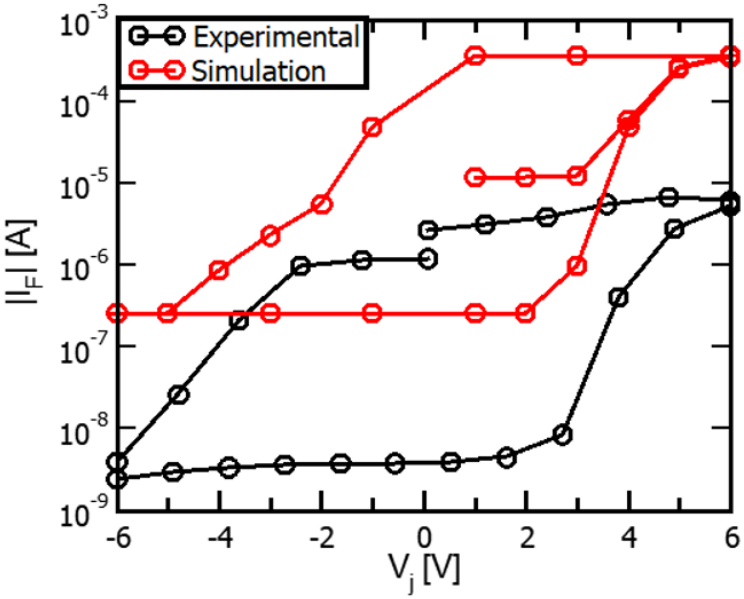


Figure 4.6: Comparison between simulated and measured resistive windows.

for each P value, once the electric field is needed the value can be retrieved with an interpolation operation on the row of the matrix that corresponds to the applied V_j; the result will be more precise the more P values are simulated and saved in the matrix. A summarization of the simulation flow is shown in Fig. 4.4.

4.3.2 Early results

As we can see from the analysis carried out about ultrathin ferroelectric BTO layers [110], the Landau parameters needed for Eq. 4.3 depends on other physical characteristic related to

Table 4.1: List of parameters used in the switching simulation, mean values taken from [110].

	μ	σ
$\overline{a}$[m/F]	4.124(T-115)·10^5	//
$\overline{b}$[m^5/C^2F]	-2.097·10^8	//
a[m/F]	-1.42·10^9	3.5·10^8
b[m^5/C^2F]	1.20·10^8	3.5·10^8
c[m^9/C^4F]	1.294·10^9	2.5·10^9
d[m^{13}/C^6F]	3.863·10^{10}	1·10^{10}
Q_{12}[m^4/C^2]	-0.043	//
s_{11}[m^2/N]	8.3·10^{-12}	//
s_{12}[m^2/N]	-2.7·10^{-12}	//
u [Δm/m]	-0.09	//
ρ [Ωm]	20-200	//
k [m^3/F]	1.44·10^{-9}	//

the lattice of the particular ferroelectric material, such that:

$$\begin{cases} a = \overline{a} - u\frac{Q_{12}}{s_{11}+s_{12}} \\ \\ b = \overline{b} + \frac{Q_{12}^2}{s_{11}+s_{12}} \end{cases} \tag{4.11}$$

where s_{11} and s_{12} are elastic compliances at constant polarization, Q_{12} is the electroresistive constant and u is the misfit strain. The investigation of such characteristics could not be performed during the Ph.D. and was beyond the scope of this thesis, therefore the values were taken from the cited paper [110]. The values that could not be retrieved from that work were chosen as fitting parameters. The misfit strain u is constant, ρ is chosen from a uniform distribution and k is a constant as in [109]. Standard deviations were chosen to be as high as possible without affecting the stability of the simulation. The reason for this choice is two-fold: the first one is to test the limits of the simulation approach against a higly non-uniform condition. The second is that such condition was more interesting in order to investigate how zones of

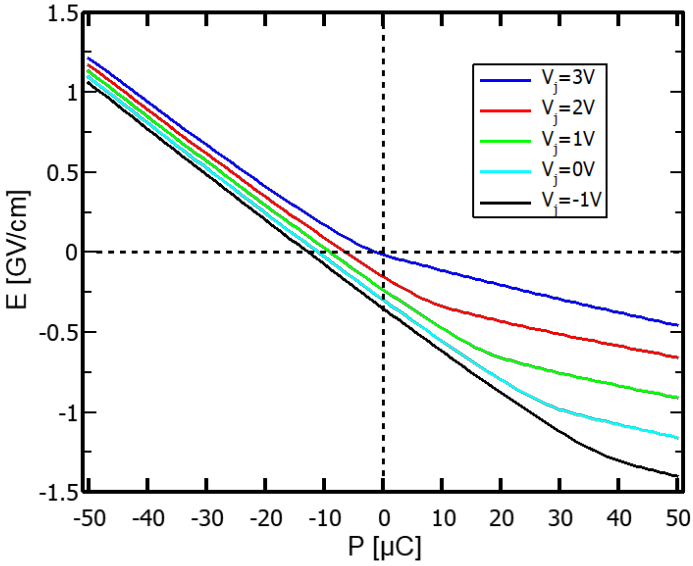


Figure 4.7: Electric fields in the ferroelectric layer as a function of polarization and applied bias.

the material with different polarization directions influenced one another.

Fig. 4.5 shows the simulated current at ±100 mV after the application of programming pulses and a pause with no bias applied both of 100 μs. The pause was much longer during the measurement, in the order of one second, but during the simulation there was no net effect after a few tens of microseconds and therefore it was chosen to limit the pause to reduce the simulation time. The total simulation time, for 22 writing pulses, with a time step Δt=40 ns, was 1.98 hours. From this picture we can pinpoint some relevant elements. First of all the magnitude of the ON/OFF ratio, about three decades, is in agreement with the one shown in Fig. 3.8, along with the presence of intermediate states reflecting the presence of different domains with different or opposite polarization on the

device area. Looking at the shape of the hysteresis it can be noted that the high current branch (full LRS) is much more shorter that the low current one, as soon as a negative bias is applied (V_j=-1 V) the current drops by a decade, while on the specular situation (full HRS) the current starts rising when V_j=3 V. This difference is mainly due to the high negative built-in field, the difference between the metal work function and the STO electron affinity is about −2 V, and of course it depends strongly on the chosen 1-D parameters. Another major issue is that all the curve is about two orders of magnitude higher with respect to the experimental one, as shown in Fig. 4.6. This latter incongruence may be solved by tweaking the Landau parameters in order to have stable polarization values of lower absolute values (e.g. rising the parameters b or c, see Fig. 4.2); but the former difference is rooted in the choice of the 1-D parameters and the absence of simmetry of the junction, looking at Fig. 3.24 one can see that the electric field remains strongly negative for both LRS and HRS. In order to take a closer look into this aspect Fig. 4.7 shows the electric fields in the ferroelectric material depending on the polarization and on the applied bias V_j. The electric fields when P=0 are always negative regardless of the applied bias, till V_j=3 V.

4.3.3 Conclusions

In this chapter, a two-dimensional ferroelectric switching simulator has been presented. First, the notion of domain and a few theories about ferroelectric switching are illustrated, in particular the Landau-Ginzburg-Devonshire theory which is exploited by the simulator. Secondly, the simulator algorithm and equations are illustrated together with simulation workarounds to reduce the simulation time. At the end of the chapter the simulation results are shown: the simulated resistive window present an ON/OFF ratio in agreement with the experimental one but the curve is two orders of magni-

tude higher and loses the LRS much faster than the measured one. The reasons for such discrepancies are probably rooted in the chosen 1-D parameters. Such values may be correct for a 1-D description of the device but are probably not suitable for a complete description that takes into account the probability of having zones with different polarization values on the device area; one could say that they are mean or effective 1-D values. Therefore, the simulator was not able to reproduce the experimental measures. The elevated number of parameters needed for both simulators, combined with the impossibility, during this research, to directly measure such parameters, set obstacles that were not overcome. In order to surpass such issues verifications with 2-D measures on the area such PFM may be useful, together with a reduction of the degrees of freedom for the 1-D simulator with the direct measurements of some of the needed parameters. Despite the problems listed above the results are very promising: the simulation times are satisfying and may be improved again, the simulator is simple enough to lend itself easily to additions and changes and it can be a powerful tool to further study the switching dynamics in FTJ devices.

Summary of results

The research activity carried out during the Ph.D. program focused on the investigation and modeling of Ferroelectric Tunnel Junctions, with emphasis on the 1-D current transport and 2-D ferroelectric switching. For this reason, the thesis has been divided into two main parts, collecting the results achieved about one-dimensional characterization and modeling in Chapter 3 and those about two-dimensional ferroelectric switching in Chapter 4.

In Chapter 3, a comprehensive investigation of the physical phenomena involved in the 1-D current transport was performed with the help of an in-house simulator written in *Mathworks MATLAB*. The reaserch started with the experimental characterization of different FTJ samples fabricated at Polifab. The measurement setup has been designed to be able to better control the devices in different situations. Two different setups has been used for: *slow measures* where relatively long time passes between the programming and the reading operation and *fast measures* where a much shorter time, in the order of 200 μs, passes between the two operations. This latter setup allowed us to characterize for the first time the retention behavior at short times after the set/reset pulses. The quick polarization loss after the set pulse is a very important feature of this device, reducing the ON/OFF ratio

by one order of magnitude. The experimental characterization was followed by a modeling activity during which a 1-D simulator for both electrostatics and current trasport has been written in MATLAB. This simulator successfully fit the aforementioned experimental measures and allowed us to study the impact of distributed tunneling, another essential feature of the device, which greatly reduces the ON/OFF ratio due to the chance for electrons to tunnel not only through the metal but also through the semiconductor electronic bands. With the aid of the simulator we explained the reason of the poor LRS retention, due to the strong negative electric field almost always present in the ferroelectric layer of the device. Moreover, such simulator can be easily extended to other device design (e.g. crested barrier FTJ) and has already been used to study the chance for very low mobility semiconductors to become bottleneck for current conduction.

In the second part, presented in Chapter 4, efforts were directed toward the development of a two-dimensional simulator for ferroelectric switching which could benefit of the 1-D simulator already written. Starting from literature results, the limitations of that approach were overcome thanks to the 1-D MATLAB code. By its exploitation, it has been possible to know the electrostatics of the device at any moment in time and therefore avoid relying on analytical estimates of certain quantities necessary for the operation of the two-dimensional simulator. The development of the switching simulator has been completed, but the results obtained were not up to the standard of the ones of 1-D simulator. An attempt has been made to reproduce the resistive hysteresis measurement, but even if the simulated ON/OFF ratio is in fairly good agreement with experiments, the simulated curve is not as symmetrical as the experimental one and it is two orders of magnitude higher than expected. The results are nevertheless promising and this simulator can undoubtedly be of help in better understanding physical phenomena involved

during ferroelectric switching in FTJs. The reason why the experimental measure could not be reproduced probably lies in the chosen 1-D and 2-D parameters, an obstacle that can be overcome with specific experimental investigations on the Landau parameters or some of the critical measurable quantities needed for the 1-D simulator.

In conclusion, the research activity carried out during the Ph.D. program contributed to the understanding and modeling of the physical phenomena involved during the FTJ current transport and retention. The developed simulators can form a basis for further analysis, especially with regard to ferroelectric switching, for which only preliminary results have been achieved.

Bibliography

[1] T. Tseng and S. Sze, "An introduction to nonvolatile memories," *Nonvolatile Memories: Materials, Devices, and Applications*, vol. 1, pp. 1–9, 2012.

[2] "Site." https://www.grandviewresearch.com/industry-analysis/the-global-semiconductor-memory, 2019.

[3] Micron Technology Inc., "2018 Analyst & Investor Event." http://investors.micron.com/, 2018.

[4] F. Masuoka, M. Asano, H. Iwahashi, T. Komuro, and S. Tanaka, "A new Flash E^2PROM cell using triple polysilicon technology," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 464–467, 1984.

[5] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells–an overview," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1248–1271, 1997.

[6] M. Lenzlinger and E. Snow, "Fowler-Nordheim tunneling into thermally grown SiO_2," *Journal of Applied physics*, vol. 40, no. 1, pp. 278–283, 1969.

[7] B. Eitan and D. Frohman-Bentchkowsky, "Hot-electron injection into the oxide in *n*-channel MOS devices," *IEEE Transactions on Electron Devices*, vol. 28, no. 3, pp. 328–340, 1981.

[8] C. Monzio Compagnoni, R. Gusmeroli, A. Sottocornola Spinelli, D. Ielmini, A. L. Lacaita, and A. Visconti, "Present status and scaling challenges for the NOR Flash memory technology," in *Solid State Electronics Research Advances* (S. Kobadze, ed.), ch. 3, pp. 101–134, Nova Science Publishers, inc, 2009.

[9] Y. Tkachev and A. Kotov, "A detailed analysis of hot-electron programming efficiency in 40 nm split-gate Flash memory cells," in *IEEE International Memory Workshop (IMW)*, pp. 1–4, 2017.

[10] C. Monzio Compagnoni, A. Goda, A. Sottocornola Spinelli, P. Feeley, A. L. Lacaita, and A. Visconti, "Reviewing the evolution of the NAND Flash technology," *Proceedings of the IEEE*, vol. 105, no. 9, pp. 1609–1633, 2017.

[11] S. Aritome, *NAND Flash memory technologies*. John Wiley & Sons, 2015.

[12] S. Lee, J.-y. Lee, I.-h. Park, J. Park, S.-w. Yun, M.-s. Kim, J.-h. Lee, M. Kim, K. Lee, T. Kim, *et al.*, "7.5 A 128 Gb 2b/cell NAND Flash memory in 14 nm technology with t_{PROG}= 640 µs and 800 MB/s I/O rate," in *IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 138–139, 2016.

[13] C. Monzio Compagnoni, R. Gusmeroli, A. Spinelli Sottocornola, A. L. Lacaita, M. Bonanomi, and A. Visconti, "Statistical model for random telegraph noise in Flash memories," *IEEE Transactions on Electron Devices*, vol. 55, no. 1, pp. 388–395, 2007.

[14] C. Monzio Compagnoni, A. Sottocornola Spinelli, R. Gusmeroli, A. L. Lacaita, S. Beltrami, A. Ghetti, and A. Visconti, "First evidence for injection statistics accuracy limitations in NAND Flash constant-current Fowler-Nordheim programming," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 165–168, IEEE, 2007.

[15] C. Monzio Compagnoni, A. Sottocornola Spinelli, R. Gusmeroli, S. Beltrami, A. Ghetti, and A. Visconti, "Ultimate accuracy for the NAND Flash program algorithm due to the electron injection statistics," *IEEE Transactions on Electron Devices*, vol. 55, no. 10, pp. 2695–2702, 2008.

[16] A. Spessot, A. Calderoni, P. Fantini, A. Sottocornola Spinelli, C. Monzio Compagnoni, F. Farina, A. L. Lacaita, and A. Marmiroli, "Variability effects on the V_T distribution of nanoscale NAND Flash memories," in *IEEE International Reliability Physics Symposium (IRPS)*, pp. 970–974, IEEE, 2010.

[17] G. Nicosia, *Performance and reliability issues of NAND Flash cells at the transition from planar to 3-D array architectures*. PhD thesis, Politecnico di Milano, Italy, 2019.

[18] N. Righetti and G. Puzzilli, "2D vs 3D NAND technology: Reliability benchmark," in *IEEE International Integrated Reliability Workshop (IIRW)*, pp. 1–6, 2017.

[19] J. A. Mandelman, R. H. Dennard, G. B. Bronner, J. K. DeBrosse, R. Divakaruni, Y. Li, and C. J. Radens, "Challenges and future directions for the scaling of dynamic random-access memory (DRAM)," *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 187–212, 2002.

[20] S. Narasimha, P. Chang, C. Ortolland, D. Fried, E. Engbrecht, K. Nummy, P. Parries, T. Ando, M. Aquilino, N. Arnold, *et al.*, "22nm high-performance SOI technology featuring dual-embedded stressors, epi-plate high-k deep-trench embedded DRAM and self-aligned via 15lm beol," in *2012 International Electron Devices Meeting*, pp. 3–3, IEEE, 2012.

[21] Y. Wang, U. Arslan, N. Bisnik, R. Brain, S. Ghosh, F. Hamzaoglu, N. Lindert, M. Meterelliyoz, J. Park, S. Tomishima, *et al.*, "Retention time optimization for eDRAM in 22nm tri-gate CMOS technology," in *2013 IEEE International Electron Devices Meeting*, pp. 9–5, IEEE, 2013.

[22] W. Mueller, G. Aichmayr, W. Bergner, E. Erben, T. Hecht, C. Kapteyn, A. Kersch, S. Kudelka, F. Lau, J. Luetzen, *et al.*, "Challenges for the DRAM cell scaling to 40nm," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, pp. 4–pp, IEEE, 2005.

[23] J. Kittl, K. Opsomer, M. Popovici, N. Menou, B. Kaczer, X. Wang, C. Adelmann, M. Pawlak, K. Tomida, A. Rothschild, *et al.*, "High-k dielectrics for future generation memory devices," *Microelectronic engineering*, vol. 86, no. 7-9, pp. 1789–1795, 2009.

[24] S.-K. Park, "Technology scaling challenge and future prospects of DRAM and NAND flash memory," in *2015 IEEE International Memory Workshop (IMW)*, pp. 1–4, IEEE, 2015.

[25] M. G. Ertosun, *Novel Capacitorless Single-Transistor DRAM Technologies*. Stanford University, 2010.

[26] M. G. Ertosun, H. Cho, P. Kapur, and K. C. Saraswat, "A nanoscale vertical double-gate single-transistor capacitorless DRAM," *IEEE electron device letters*, vol. 29, no. 6, pp. 615–617, 2008.

[27] H.-J. Cho, F. Nemati, R. Roy, R. Gupta, K. Yang, M. Ershov, S. Banna, M. Tarabbia, C. Salling, D. Hayes, *et al.*, "A novel capacitor-less DRAM cell using thin capacitively-coupled thyristor (tcct)," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, pp. 311–314, IEEE, 2005.

[28] G. J. Myers, "Advances in computer architecture," tech. rep., 1978.

[29] S. Borkar and A. A. Chien, "The future of microprocessors," *Communications of the ACM*, vol. 54, no. 5, pp. 67–77, 2011.

[30] B. Jacob, D. Wang, and S. Ng, *Memory systems: cache, DRAM, disk*. Morgan Kaufmann, 2010.

[31] J. Backus, "Can programming be liberated from the von Neumann style? a functional style and its algebra of programs," *Communications of the ACM*, vol. 21, no. 8, pp. 613–641, 1978.

[32] K. Itoh, *VLSI memory chip design*, vol. 5. Springer Science & Business Media, 2013.

[33] Y. Morita, H. Fujiwara, H. Noguchi, Y. Iguchi, K. Nii, H. Kawaguchi, and M. Yoshimoto, "Area optimization in 6T and 8T SRAM cells considering V_{th} variation in future processes," *IEICE transactions on electronics*, vol. 90, no. 10, pp. 1949–1956, 2007.

[34] H.-S. P. Wong and S. Salahuddin, "Memory leads the way to better computing," *Nature nanotechnology*, vol. 10, no. 3, pp. 191–194, 2015.

[35] G. W. Burr, B. N. Kurdi, J. C. Scott, C. H. Lam, K. Gopalakrishnan, and R. S. Shenoy, "Overview of candidate device technologies for storage-class memory," *IBM Journal of Research and Development*, vol. 52, no. 4.5, pp. 449–464, 2008.

[36] S. Lai and T. Lowrey, "OUM-A 180 nm nonvolatile memory cell element technology for stand alone and embedded applications," in *International Electron Devices Meeting. Technical Digest (Cat. No. 01CH37224)*, pp. 36–5, IEEE, 2001.

[37] C. Chappert, A. Fert, and F. N. Van Dau, "The emergence of spin electronics in data storage," *Nanoscience And Technology: A Collection of Reviews from Nature Journals*, pp. 147–157, 2010.

[38] J. Slaughter, N. Rizzo, F. Mancoff, R. Whig, K. Smith, S. Aggarwal, and S. Tehrani, "Toggle and spin-torque MRAM: Status and outlook," *Journal of the Magnetic Society of Japan*, vol. 5, no. 1, pp. 171–176, 2010.

[39] J. C. Slonczewski, "Current-driven excitation of magnetic multilayers," *Journal of Magnetism and Magnetic Materials*, vol. 159, no. 1-2, pp. L1–L7, 1996.

[40] J. J. Nowak, R. P. Robertazzi, J. Z. Sun, G. Hu, J.-H. Park, J. Lee, A. J. Annunziata, G. P. Lauer, R. Kothandaraman, E. J. O'Sullivan, *et al.*, "Dependence of voltage and size on write error rates in spin-transfer torque magnetic random-access memory," *IEEE Magnetics Letters*, vol. 7, pp. 1–4, 2016.

[41] R. Carboni, S. Ambrogio, W. Chen, M. Siddik, J. Harms, A. Lyle, W. Kula, G. Sandhu, and D. Ielmini, "Modeling of breakdown-limited endurance in spin-transfer torque magnetic memory under pulsed cycling regime," *IEEE Transactions on Electron Devices*, vol. 65, no. 6, pp. 2470–2478, 2018.

[42] D. Apalkov, B. Dieny, and J. Slaughter, "Magnetoresistive random access memory," *Proceedings of the IEEE*, vol. 104, no. 10, pp. 1796–1830, 2016.

[43] L. Landau, "The theory of phase transitions," *Nature*, vol. 138, no. 3498, pp. 840–841, 1936.

[44] T. Böscke, J. Müller, D. Bräuhaus, U. Schröder, and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Applied Physics Letters*, vol. 99, no. 10, p. 102903, 2011.

[45] J. Müller, T. Böscke, S. Müller, E. Yurchuk, P. Polakowski, J. Paul, D. Martin, T. Schenk, K. Khullar, A. Kersch, *et al.*, "Ferroelectric hafnium oxide: A CMOS-compatible and highly scalable approach to future ferroelectric memories," in *2013 IEEE International Electron Devices Meeting*, pp. 10–8, IEEE, 2013.

[46] M. H. Park, Y. H. Lee, T. Mikolajick, U. Schroeder, and C. S. Hwang, "Review and perspective on ferroelectric HfO2-based thin films for memory applications," *MRS Communications*, vol. 8, no. 3, pp. 795–808, 2018.

[47] V. K. M. V. M. Kondo, V. S. K. Singh, and V. H. Ishiwara, "New ferroelectric material for embedded fram lsis," *Fujitsu Sci. Tech. J*, vol. 43, no. 4, pp. 502–507, 2007.

[48] J. Müller, P. Polakowski, S. Mueller, and T. Mikolajick, "Ferroelectric hafnium oxide based materials and devices: Assessment of current status and future prospects," *ECS Journal of Solid State Science and Technology*, vol. 4, no. 5, p. N30, 2015.

[49] Z. Wen and D. Wu, "Ferroelectric tunnel junctions: Modulations on the potential barrier," *Advanced Materials*, vol. 32, no. 27, p. 1904123, 2020.

[50] a. L. Esaki, R. Laibowitz, and P. Stiles, "Polar switch," *IBM Tech. Discl. Bull*, vol. 13, no. 2161, p. 114, 1971.

[51] B. Max, M. Hoffmann, S. Slesazeck, and T. Mikolajick, "Ferroelectric tunnel junctions based on ferroelectric-dielectric $Hf_{0.5}Zr_{0.5}O2$/A12O3 capacitor stacks," in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, pp. 142–145, IEEE, 2018.

[52] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics*. Cambridge University Press, 2018.

[53] Y. Qi and A. M. Rappe, "Designing ferroelectric field-effect transistors based on the polarization-rotation effect for low operating voltage and fast switching," *Physical Review Applied*, vol. 4, no. 4, p. 044014, 2015.

[54] A. Chanthbouala, A. Crassous, V. Garcia, K. Bouzehouane, S. Fusil, X. Moya, J. Allibe, B. Dlubak, J. Grollier, S. Xavier, *et al.*, "Solid-state memories based on ferroelectric tunnel junctions," *Nature nanotechnology*, vol. 7, no. 2, p. 101, 2012.

[55] Z. Wen, C. Li, D. Wu, A. Li, and N. Ming, "Ferroelectric-field-effect-enhanced electroresistance in metal/ferroelectric/semiconductor tunnel junctions," *Nature materials*, vol. 12, no. 7, pp. 617–621, 2013.

[56] J. Li, N. Li, C. Ge, H. Huang, Y. Sun, P. Gao, M. He, C. Wang, G. Yang, and K. Jin, "Giant electroresistance in ferroionic tunnel junctions," *IScience*, vol. 16, pp. 368–377, 2019.

[57] G. Radaelli, D. Gutiérrez, M. Qian, I. Fina, F. Sánchez, L. Baldrati, J. Heidler, C. Piamonteze, R. Bertacco, and J. Fontcuberta, "Strain-controlled responsiveness of slave half-doped manganite $La_{0.5}Sr_{0.5}MnO3$ layers inserted in BaTiO3 ferroelectric tunnel junctions," *Advanced Electronic Materials*, vol. 2, no. 12, p. 1600368, 2016.

[58] Z. Xi, J. Ruan, C. Li, C. Zheng, Z. Wen, J. Dai, A. Li, and D. Wu, "Giant tunnelling electroresistance in metal/ferroelectric/semiconductor tunnel junctions by engineering the Schottky barrier," *Nature communications*, vol. 8, no. 1, pp. 1–9, 2017.

[59] "Polifab," 2021.

[60] S. Suzuki, T. Yamamoto, H. Suzuki, K. Kawaguchi, K. Takahashi, and Y. Yoshisato, "Fabrication and characterization of $Ba_{1-x}K_xBiO3$/Nb-doped SrTiO3 all-oxide-type Schottky junctions," *Journal of applied physics*, vol. 81, no. 10, pp. 6830–6836, 1997.

[61] H. Yamane, H. Masumoto, T. Hirai, H. Iwasaki, K. Watanabe, N. Kobayashi, Y. Muto, and H. Kurosawa, "Y-Ba-Cu-O superconducting films prepared on SrTiO3 substrates by chemical vapor deposition," *Applied physics letters*, vol. 53, no. 16, pp. 1548–1550, 1988.

[62] C. Marchiori, M. Sousa, A. Guiller, H. Siegwart, J.-P. Locquet, J. Fompeyrine, G. Norga, and J. W. Seo, "Thermal stability of the SrTiO3/(Ba, Sr)O stacks epitaxially grown on Si," *Applied physics letters*, vol. 88, no. 7, p. 072913, 2006.

[63] M. Asa, L. Baldrati, C. Rinaldi, S. Bertoli, G. Radaelli, M. Cantoni, and R. Bertacco, "Electric field control of magnetic properties and electron transport in BaTiO3-based multiferroic heterostructures," *Journal of Physics: Condensed Matter*, vol. 27, no. 50, p. 504004, 2015.

[64] G. Radaelli, S. Brivio, I. Fina, and R. Bertacco, "Correlation between growth dynamics and dielectric properties of epitaxial BaTiO3 films," *Applied Physics Letters*, vol. 100, no. 10, p. 102904, 2012.

[65] M. Asa and R. Bertacco, "Impact of semiconducting electrodes on the electroresistance of ferroelectric tunnel junctions," *Applied Physics Letters*, vol. 112, no. 8, p. 082903, 2018.

[66] "Manual." https://res.cloudinary.com/iwh/image/upload/q_auto,g_center/assets/1/26/Documents/Agilent/34970A/34970a_doc_1.pdf, 2019.

[67] "Manual." https://usermanual.wiki/Document/A4140B.462139795/html, 2019.

[68] "Manual." hhttps://assets.tequipment.net/assets/1/26/TTi_TGA12100_Series_Arbitrary_Generator_Instruction_Manual.pdf, 2019.

[69] "Manual." https://neurophysics.ucsd.edu/Manuals/Tektronix/TDS%20200_Digital%20Real-Time%200scilloscope.pdf, 2019.

[70] A. Quindeau, D. Hesse, and M. Alexe, "Programmable ferroelectric tunnel memristor," *Frontiers in Physics*, vol. 2, p. 7, 2014.

[71] Z. Wen, D. Wu, and A. Li, "Memristive behaviors in Pt/BaTiO3/Nb: Srtio3 ferroelectric tunnel junctions," *Applied Physics Letters*, vol. 105, no. 5, p. 052910, 2014.

[72] M. J. Highland, T. T. Fister, M.-I. Richard, D. D. Fong, P. H. Fuoss, C. Thompson, J. A. Eastman, S. K. Streiffer, and G. B. Stephenson, "Polarization switching without domain formation at the intrinsic coercive field in ultrathin ferroelectric PbTiO3," *Physical review letters*, vol. 105, no. 16, p. 167601, 2010.

[73] "Datasheet." https://www.analog.com/en/products/adg1402.html, 2020.

[74] "Datasheet." https://www.analog.com/en/products/adg1221.html, 2020.

[75] S.-C. Chang, A. Naeemi, D. E. Nikonov, and A. Gruverman, "Theoretical approach to electroresistance in ferroelectric tunnel junctions," *Physical Review Applied*, vol. 7, no. 2, p. 024005, 2017.

[76] M. Y. Zhuravlev, R. F. Sabirianov, S. Jaswal, and E. Y. Tsymbal, "Giant electroresistance in ferroelectric tunnel junctions," *Physical Review Letters*, vol. 94, no. 24, p. 246802, 2005.

[77] M. Ieong, P. M. Solomon, S. Laux, H.-S. Wong, and D. Chidambarrao, "Comparison of raised and Schottky source/drain MOSFETS using a novel tunneling contact model," in *International Electron Devices Meeting 1998. Technical Digest (Cat. No. 98CH36217)*, pp. 733–736, IEEE, 1998.

[78] K. Matsuzawa, K. Uchida, and A. Nishiyama, "A unified simulation of Schottky and ohmic contacts," *IEEE Transactions on Electron Devices*, vol. 47, no. 1, pp. 103–108, 2000.

[79] K. Matsuzawa, "Impact of random dopant fluctuation on size-dependence of contact resistance," *Japanese Journal of Applied Physics*, vol. 50, no. 12R, p. 124302, 2011.

[80] P. Palestri, N. Barin, D. Brunel, C. Busseret, A. Campera, P. A. Childs, F. Driussi, C. Fiegna, G. Fiori, R. Gusmeroli, G. Iannaccone, M. Karner, H. Kosina, A. L. Lacaita, E. Langer, B. Majkusiak, C. M. Compagnoni, A. Poncet, E. Sangiorgi, L. Selmi, A. S. Spinelli, and J. Walczak, "Comparison of modeling approaches for the capacitance–voltage and current–voltage characteristics of advanced gate stacks," *IEEE Transactions on Electron Devices*, vol. 54, no. 1, pp. 106–114, 2007.

[81] H. Levy and F. Lessman, *Finite difference equations*. Courier Corporation, 1992.

[82] "Wikipedia." `https://en.wikipedia.org/wiki/Newton%27s_method`, 2020.

[83] O. Simonetti, T. Maurel, and M. Jourdain, "Characterization of ultrathin metal–oxide–semiconductor structures using coupled current and capacitance–voltage models based on quantum calculation," *Journal of applied physics*, vol. 92, no. 8, pp. 4449–4458, 2002.

[84] E. Nadimi, C. Golz, M. Trentzsch, L. Herrmann, K. Wieczorek, and C. Radehaus, "Tunneling effective mass of electrons in lightly n-doped SiO_xN_y gate insulators," *IEEE Transactions on Electron Devices*, vol. 55, no. 9, pp. 2462–2468, 2008.

[85] B. Max, T. Mikolajick, M. Hoffmann, S. Slesazeck, and T. Mikolajick, "Retention characteristics of $Hf_{0.5}Zr_{0.5}O2$-based ferroelectric tunnel junctions," in *2019 IEEE 11th International Memory Workshop (IMW)*, pp. 1–4, 2019.

[86] N. Gong and T.-P. Ma, "Why is Fe–HfO2 more suitable than PZT or SBT for scaled nonvolatile 1-T memory cell? a retention perspective," *IEEE Electron Device Letters*, vol. 37, no. 9, pp. 1123–1126, 2016.

[87] T. Ma and J.-P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," *IEEE Electron Device Letters*, vol. 23, no. 7, pp. 386–388, 2002.

[88] J. Ravichandran, W. Siemons, M. Scullin, S. Mukerjee, M. Huijben, J. Moore, A. Majumdar, and R. Ramesh, "Tuning the electronic effective mass in double-doped SrTiO3," *Physical Review B*, vol. 83, no. 3, p. 035101, 2011.

[89] H. Kohlstedt, N. Pertsev, J. R. Contreras, and R. Waser, "Theoretical current-voltage characteristics of ferroelectric tunnel junctions," *Physical Review B*, vol. 72, no. 12, p. 125341, 2005.

[90] M. Kobayashi, Y. Tagawa, F. Mo, T. Saraya, and T. Hiramoto, "Ferroelectric HfO2 tunnel junction memory with high ter and multi-level operation featuring metal replacement process," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 134–139, 2019.

[91] D. Pantel and M. Alexe, "Electroresistance effects in ferroelectric tunnel barriers," *Physical Review B*, vol. 82, no. 13, p. 134105, 2010.

[92] F. Mo, Y. Tagawa, T. Saraya, T. Hiramoto, and M. Kobayashi, "Scalability study on ferroelectric-HfO2 tunnel junction memory based on non-equilibrium green function method with self-consistent potential," in *2018 IEEE International Electron Devices Meeting (IEDM)*, pp. 16.3.1–16.3.4, 2018.

[93] F. Mo, Y. Tagawa, T. Saraya, T. Hiramoto, and M. Kobayashi, "Scalability study on ferroelectric-HfO2 tunnel junction memory based on non-equilibrium green function method," in *2019 19th Non-Volatile Memory Technology Symposium (NVMTS)*, pp. 1–5, 2019.

[94] M. Hoffmann, M. Pešić, S. Slesazeck, U. Schroeder, and T. Mikolajick, "On the stabilization of ferroelectric negative capacitance in nanoscale devices," *Nanoscale*, vol. 10, no. 23, pp. 10891–10899, 2018.

[95] T. Rollo, F. Blanchini, G. Giordano, R. Specogna, and D. Esseni, "Stabilization of negative capacitance in ferroelectric capacitors with and without a metal interlayer," *Nanoscale*, vol. 12, no. 10, pp. 6121–6129, 2020.

[96] Y. González, A. Hadj Youssef, A. Dörfler, R. Katoch, A. El Mesoudy, A. Sarkissian, D. Drouin, and A. Ruediger, "Competing tunneling conduction mechanisms in oxygen deficient $Hf_{0.5}Zr_{0.5}O2$," *Applied Physics Letters*, vol. 119, no. 13, p. 132906, 2021.

[97] N. Andreeva, A. Petrov, A. Petraru, A. Petukhov, and A. Rybkin, "Mechanisms of electron transport in BaTiO3 ultrathin epitaxial films in the temperature range 40K-295K," *Materials Research Express*, vol. 6, no. 2, p. 026427, 2018.

[98] X. Liu, J. Zheng, D. Wang, P. Musavigharavi, E. A. Stach, R. Olsson III, and D. Jariwala, "A CMOS compatible aluminum scandium nitride-based ferroelectric tunnel junction memristor," *arXiv preprint arXiv:2012.10019*, 2020.

[99] K. Ota, M. Yamaguchi, R. Berdan, T. Marukame, Y. Nishi, K. Matsuo, K. Takahashi, Y. Kamiya, S. Miyano, J. Deguchi, *et al.*, "Performance maximization of in-memory reinforcement learning with variability-controlled $Hf_{1-x}Zr_xO2$ ferroelectric tunnel junctions," in *2019 IEEE International Electron Devices Meeting (IEDM)*, pp. 6–2, IEEE, 2019.

[100] A. N. KOLMOGOROV *Izv. Akad. Nauk. USSR Ser. Math.*, vol. 3, p. 355, 1937.

[101] M. Avrami, "Granulation, phase change, and microstructure kinetics of phase change. iii," *The Journal of chemical physics*, vol. 9, no. 2, pp. 177–184, 1941.

[102] K. J. Choi, M. Biegalski, Y. Li, A. Sharan, J. Schubert, R. Uecker, P. Reiche, Y. Chen, X. Pan, V. Gopalan, *et al.*, "Enhancement of ferroelectricity in strained BaTiO3 thin films," *Science*, vol. 306, no. 5698, pp. 1005–1009, 2004.

[103] N. Pertsev, A. Zembilgotov, and A. Tagantsev, "Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films," *Physical review letters*, vol. 80, no. 9, p. 1988, 1998.

[104] J. Jo, H. Han, J.-G. Yoon, T. Song, S.-H. Kim, and T. Noh, "Domain switching kinetics in disordered ferroelectric thin films," *Physical review letters*, vol. 99, no. 26, p. 267602, 2007.

[105] A. Gruverman, B. J. Rodriguez, C. Dehoff, J. Waldrep, A. Kingon, R. Nemanich, and J. Cross, "Direct studies of domain switching dynamics in thin film ferroelectric capacitors," *Applied Physics Letters*, vol. 87, no. 8, p. 082902, 2005.

[106] A. K. Tagantsev, I. Stolichnov, N. Setter, J. S. Cross, and M. Tsukada, "Non-Kolmogorov-Avrami switching kinetics in ferroelectric thin films," *Physical Review B*, vol. 66, no. 21, p. 214109, 2002.

[107] L. Landau and I. Khalatnikov, "On the anomalous absorption of a sound near to points of phase transition of the second kind," in *Dokl. Akad. Nauk SSSR*, vol. 96, p. 469, 1954.

[108] L. D. Landau and V. L. Ginzburg, "On the theory of superconductivity," *Zh. Eksp. Teor. Fiz.*, vol. 20, p. 1064, 1950.

[109] M. Hoffmann, A. I. Khan, C. Serrao, Z. Lu, S. Salahuddin, M. Pešić, S. Slesazeck, U. Schroeder, and T. Mikolajick, "Ferroelectric negative capacitance domain dynamics," *Journal of Applied Physics*, vol. 123, no. 18, p. 184101, 2018.

[110] W. Ma, J. Zhong, and W. Jian, "Curie phase transition and critical size for ferroelectricity in strained ultrathin PbTiO3 and BaTiO3: A phenomenological study," *Ferroelectrics*, vol. 507, no. 1, pp. 86–101, 2017.

List of publications

International journals

1. **G. Franchini**, G. Malavena, C. Monzio Compagnoni, and A. Sottocornola Spinelli, "Investigation of the Meyer-Neldel rule in MOSFETs," *IEEE Electron Device Letters*, vol. 41, no. 12, pp. 1821–1824, 2020.
2. **G. Franchini**, A. Sottocornola Spinelli, G. Nicosia, I. Fumagalli, M. Asa, C. Groppi, C. Rinaldi, R. Bertacco C. Monzio Compagnoni, "Characterization and Modeling of Current Transport in Metal/Ferroelectric/Semiconductor Tunnel Junctions," *IEEE Transactions on Electron Devices*, vol. 67, no. 9, pp. 3729-3735, Sept. 2020.