

# Redox based Non-volatile Memory Devices

Vikas Rana, Susanne Hoffmann-Eifert and Rainer Waser

Peter Grünberg Institute PGI-7 – Electronic Materials

Research Center Juelich, 52425, Germany

## Contents

1. Introduction.....	2
2. Resistive Switching.....	2
3. Processes during Redox based Resistive Switching.....	3
4. Types of Redox based Resistive Switching.....	4
5. Fabrication of HfO <sub>2</sub> based ReRAM Devices.....	5
(a) <i>Thin Film Plasma assisted Atomic Layer Deposition</i> .....	6
(b) <i>Photolithography</i> .....	8
(c) <i>Sputter Deposition of Metal Films</i> .....	11
(d) <i>Electrical Characterization of HfO<sub>2</sub> based ReRAM Devices</i> .....	12
6. Conclusion.....	14
References .....	14

## 1. Introduction

Research in the field of random access memories (RAM) is concerned with three cardinal issues, true non-volatility (NV), speed, and extended scaling. The advent of a non-volatile memory with a high (Gbit) integration density would merge the advantages of today's DRAM and mass storage devices, while stripping off the drawbacks. Especially as embedded memory, this would represent another revolutionary step in information technology, specifically in mobile applications and ubiquitous computing.

An ideal non-volatile memory comprises write and (preferably non-destructive) read operations at speeds comparable to those of logic devices, low energy consumption, infinite retention, and an infinite number of read and write cycles. A variety of physical principles have already been investigated in order to advance towards this ideal memory. A promising candidate for future non-volatile memories is Resistive RAM (ReRAM), based on materials, which show a significant change in the resistance upon application of a voltage bias above a critical threshold and a switch back if polarized with the reverse voltage. The insulators are typically binary ( $\text{HfO}_2$ ,  $\text{Ta}_2\text{O}_5$ ,  $\text{TiO}_2$ ,  $\text{NiO}_x$ ,  $\text{WO}_3$ ), or ternary transition metal oxides, doped with other transition metal oxides, ( $\text{SrTiO}_3$  doped with Mn, Cr, Fe). Although these materials show promising properties, the involved switching mechanisms are still content of current research activities and the integration of these materials in the ReRAM architectures is in the early stage of development.

## 2. Resistive Switching

Resistive switching is the physical phenomenon that consists on the sudden and non-volatile change of the resistance due to the application of electric stress, typically voltage or current pulsing. This effect may allow the fabrication of future novel electronic memory concepts, such as non-volatile random access memories (RAM), hence, it is also termed resistive RAM, RRAM, or ReRAM. Generally, the resistive switching phenomena in TMO can be roughly classified in **unipolar** or **bipolar** memory effect. In the unipolar case, the memory state of the system can be switched by successive application of electric stress of either the same or opposite polarities. In contrast, the bipolar memories can be toggled between the memory states by application of successive electric stress of alternate polarity.

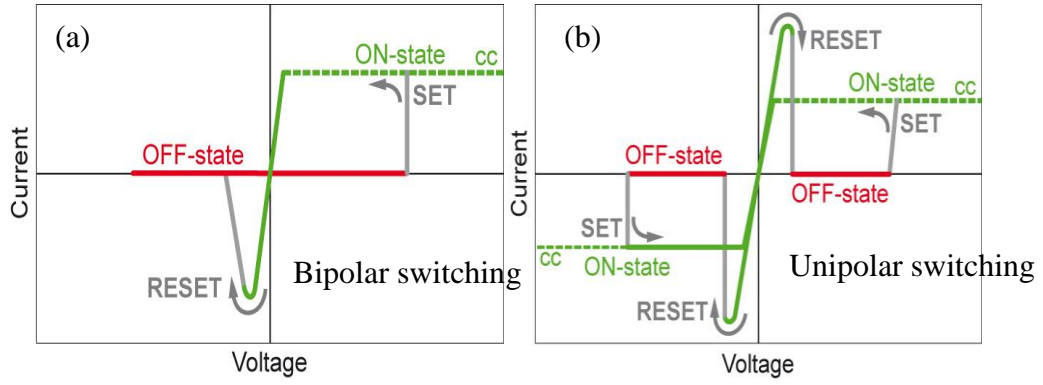


Figure 1: A demonstrative example of Current–Voltage characteristic of (a) bipolar resistive switching (b) unipolar resistive switching in transition metal oxide thin film.

The physical origin of the mechanism behind the resistive switching effect includes Schottky barriers, interfaces and electric 'faucets', spatial inhomogeneity, trapping of charge carriers, oxygen vacancy migration, filamentary path formation, correlation effects and Mott metal-insulator transitions, etc.

### 3. Processes during Redox based Resistive Switching

All conceivable processes [3] which are relevant for the electroforming and the switching process of ReRAM cells are schematically sketched in Figure 2. The metal electrodes  $M'$  and  $M''$  only carry electronic current, whereas the resistive material  $I$ , which is denoted as an ionic solid  $MX$  here, may carry electronic and ionic currents. Given the current direction in Figure 2, the ionic current may consist of anions  $X^-$  and of cations  $M^+$  moving to the left and to the right, respectively. In addition, ions  $M'^+$  may stem from the anode metal. The relative current contributions strongly depend on the type of ReRAM cells and the operation conditions. Joule heating will typically occur in the interior of the  $MX$  layer and/or close to a contact. The ionic partial current in the  $MX$  leads to electrochemical reactions, oxidation at the anode and reduction at the cathode. The specific electrochemical interface reaction is determined by the type of ReRAM cell. In addition, the ionic current may be, at least partially, blocked at the electrode interfaces. This leads to a so-called concentration polarization, i.e. an accumulation of the mobile ions near one electrode and a depletion layer near the other. Except in the (typically very narrow) space charge regions, this process is compensated by local redox reactions, i.e. a change in the average valence of the counter ions of the  $MX$  phase. As another result of the concentration polarization phase transformations and space charges can occur. Phase transformation may, of course, also result from the temperature increase due to Joule heating, and from electrochemical interface reactions. Furthermore, ad-atom diffusion may take place at interfaces. The main driving forces for the processes just described can be summarized as

- Gradients of the electrochemical potential  $\nabla\eta$
- Temperature gradients  $\nabla T$

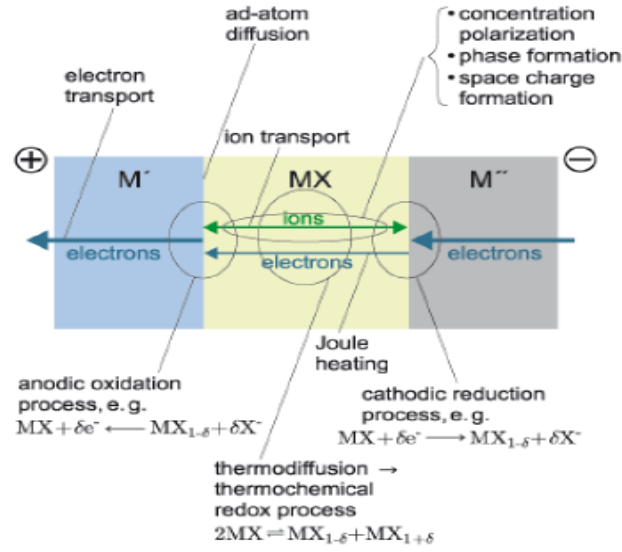


Figure 2: Overview of all processes that may be relevant in the resistive switching.  $M'$  and  $M''$  denote the electrodes, whereas  $MX$  is a mixed ionic electronics conducting component.

For the stability of states (ON and OFF) and the long term reliability, one should keep in mind that only *one* of the states, ON-state or OFF-state or any intermediate state, can be **thermodynamically stable**. Due to the nature of redox-based resistive switching, the other state(s) must be **metastable**. They are frozen-in after a kinetically fast switching event. The reason for only *one* state being the thermodynamically stable state is due to the fact that there may be only *one* arrangement of ions and atoms which has the lowest free energy. This is different to ferroelectric and ferromagnetic systems in which the states with opposite polarization direction may both thermodynamically stable.

#### 4. Types of Redox based Resistive Switching

A schematic classification of resistive switching [4] is shown in Figure 3. Depending on ion migration and coupled redox processes in the MIM cell [3], the ReRAMs can be classified into three categories: the bipolar **electrochemical metallization** memory effect (**ECM**) relies on an electrochemically active electrode metal such as Ag, the drift of the highly mobile  $Ag^+$  cations in the ion conducting I-layer, their discharge at the (inert) counter electrode leading to a growth of Ag dendrites which form a highly conductive filament in the ON-state of the cell. Upon reversal of polarity of the applied voltage, an electrochemical dissolution of these filaments takes place, resetting the system into the OFF-state. Secondly, the **valence change memory** effect (**VCM**) occurs in a wide range of metal oxides and is (typically) triggered by a migration

of anions, such as oxygen anions (which are usually described by the motion of the corresponding vacancies, i.e. oxygen vacancies).

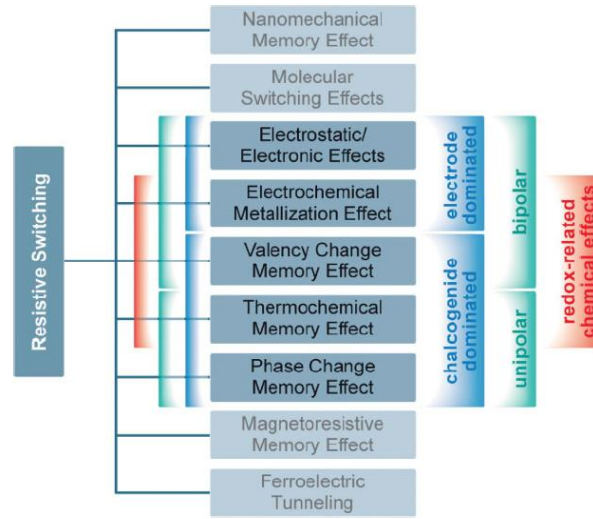


Figure 3. Classification of the resistive switching effects, which are considered for non-volatile memory applications. The switching mechanisms based on thermal, chemical, and electronic/electrostatic effects [4].

A subsequent change of the stoichiometry leads to a redox reaction expressed by a valence change of the cation sublattice and a change in the electronic conductivity in front of an electrode interface. This bipolar memory switching is induced by voltage pulses, where the polarity of the pulse determines the direction of the change, i.e. reduction or oxidation. A third type relies on a **thermochemical memory effect (TCM)** due to a current-induced increase of the temperature which leads to a redox-related change of the stoichiometry along a discharge filament, and a subsequent freezing-in of this ON-state. A differently shaped current pulse disrupts the conductive filament again to return the cell into the OFF-state.

## 5. Fabrication of HfO<sub>2</sub> based ReRAM Devices

This manuscript will elucidate the valence change memory effect in oxide materials and clarify the suitability of these materials in future RRAM. In this section, we will describe the fabrication and characterization process of HfO<sub>2</sub> based ReRAM cells. A simple schematic diagram and SEM micrograph of the ReRAM cell are shown in Figure 4.

The fabrication of the HfO<sub>2</sub> based ReRAM devices can be roughly divided into the following steps:

- 1) Starting with an oxidized silicon substrate with structured Pt bottom electrodes
- 2) Deposition of the oxide thin film (e.g. HfO<sub>2</sub>) by (Plasma Enhanced) Atomic Layer Deposition (PE-ALD)

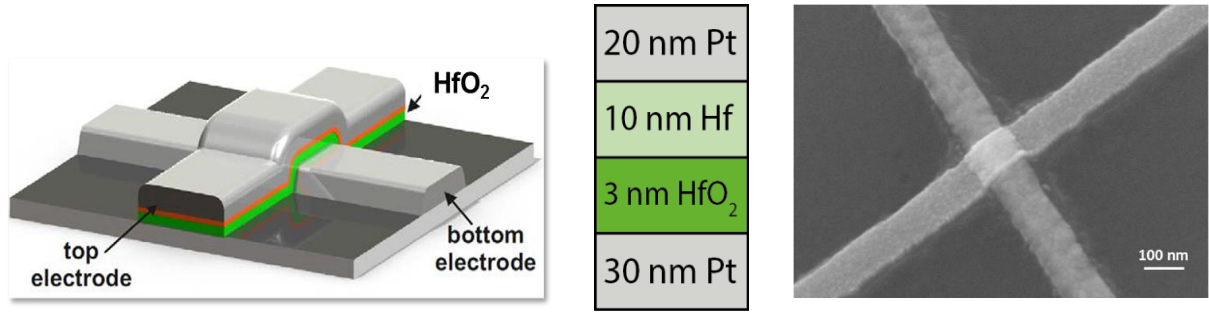


Figure 4: Simple crossbar structure for electrical testing (schematic, stack sequence and top view under electron microscope).

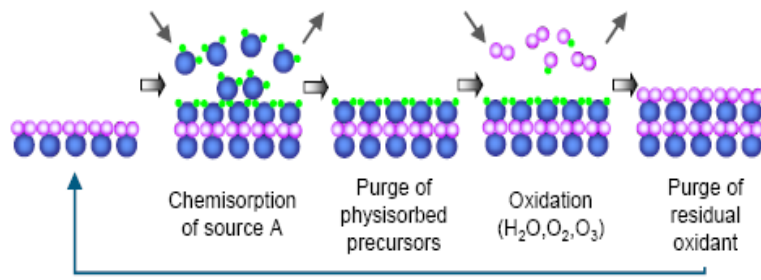
- 3) Fabrication of a resist mask by photolithography which has to be aligned to the bottom electrode structure
- 4) Metallization and structuring of the top electrode by means of a lift-off technique.

Sputtering, pulsed laser deposition (PLD), atomic layer deposition (ALD) and thermal evaporation are widely used methods for oxide and metal thin film deposition. In this manuscript, the metal thin film for the electrode formation is deposited by the sputtering process for the bottom electrode and by the thermal evaporation process for the top electrode. The deposition of the switching material, i.e. thin oxide film, is restricted to only the ALD process.

#### (a) Thin Film Deposition by ALD Method

Atomic layer deposition (ALD) is a gas phase method based on alternate saturating surface reactions and the film growth is thereby self-limiting. In contrast to Chemical Vapor Deposition (CVD) where the precursor may react in the gas phase and is decomposed on the reaction surface, in the ALD process the precursor is chemisorbed at surface sites and any adsorbed molecules are removed in a purge process step. In the following step surface ligands of the chemisorbed precursor molecules are exchanged against  $\text{OH}^-$  groups by offering a reactive gas. Therefore, the ALD is a surface-reaction controlled process and thereby allows homogeneity and excellent film thickness controllability to be achieved on complex structures. The ALD is one of the most important technologies in today's MOS gate oxide growth, DRAM and FLASH memory production. A comprehensive description of all details of the ALD technique and the involved chemical mechanisms can be found in Ref. [1].

The different stages of one ALD cycle are displayed in Figure 5:



- 1) Chemisorption of precursor molecules at surface sites;
- 2) Purge of physisorbed precursor molecules;
- 3) Oxidation of the chemisorbed species;
- 4) Purge of residual oxidant.

Figure 5: Principle of Atomic Layer deposition (ALD).

A schematic of the process flow for growth of multi component oxides is depicted in Fig. 6a. If all process steps are executed properly, the growth rate of the films saturates at a value which corresponds to the available surface sites, which means that any increase in precursor supply should not further increase the growth rate. The value depends besides others on the type of the substrate, the size and structure of the precursor molecule, and slightly on the temperature. The saturation behavior of the growth rate is shown in Fig. 6b.

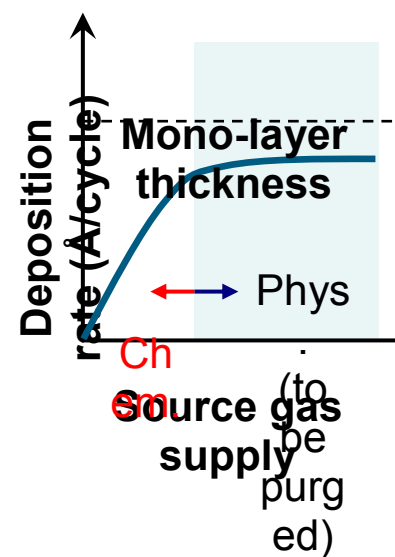
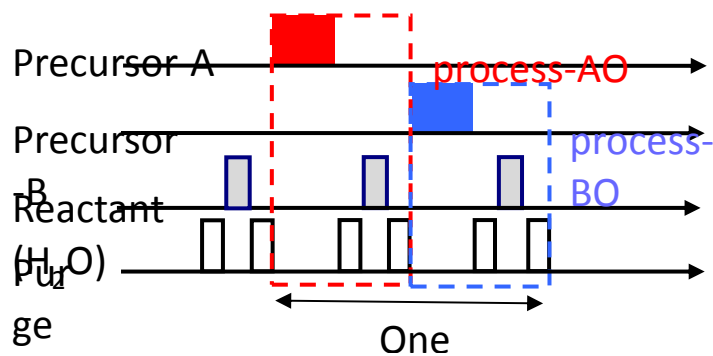


Figure 6: a) ALD process flow for multi component oxides, and b) a typical saturation dependence.

For substrate purposes, commercially available (100) oriented silicon wafers with a layer of about 400 nm-thick thermal SiO<sub>2</sub> will be used cut into pieces of 2 x 2 cm<sup>2</sup>. The Ti/Pt bottom electrodes were deposited by sputtering and were structured by means of photolithography and reactive ion etching. The plasma enhanced ALD of the thin HfO<sub>2</sub> layer is performed in an FlexAl<sup>TM</sup> remote plasma ALD system of Oxford Plasma Technology Inc., which is attached to

the Multi Cluster System in the Helmholtz Nanofacility (HNF). A schematic of the FlexAl<sup>TM</sup>-reactor and image of the FlexAl<sup>®</sup> tool are shown in Fig. 7. A dummy run for cleaning lines and reactor has been done prior to the deposition. The substrate is put on the susceptor disc and loaded to the reactor through the load lock, which is evacuated to low pressure of about  $10^{-7}$  mbar. Argon is chosen as the carrier and the purge gas. The substrate heater is set to 300°C. The Tetra(ethylmethyl)amino-hafnium precursor container is heated to 70°C and the precursor is fed via bubbling with the Ar-flow of about 250 sccm. For stabilization of the reactor conditions, 3 pre-conditioning steps (i.e. pumping to base pressure, heating of table at higher Ar-pressure of 200 mbar, stabilizing the pressure for processing 80 mTorr at 500 sccm) are done prior to the repetition of the ALD cycles. One ALD cycle for the HfO<sub>2</sub> growth from TEMA-Hf precursor and oxygen plasma consists basically of 6 steps: precursor dose 1 s / precursor purge 15 s / pumping 3 s / plasma gas stabilization 3 s / O<sub>2</sub> plasma 3 s / post plasma purge 6 s. This process is repeated until the desired film thickness is achieved. The growth rate for this process was characterized before as 0.1 nm / cycle. Due to the handling system, the wafers can be loaded and unloaded while the substrate stage stays at 300°C. All ALD valves should be switched OFF before the system is opened against the transfer. The base pressure of the system is in the range of  $10^{-6}$  mbar.

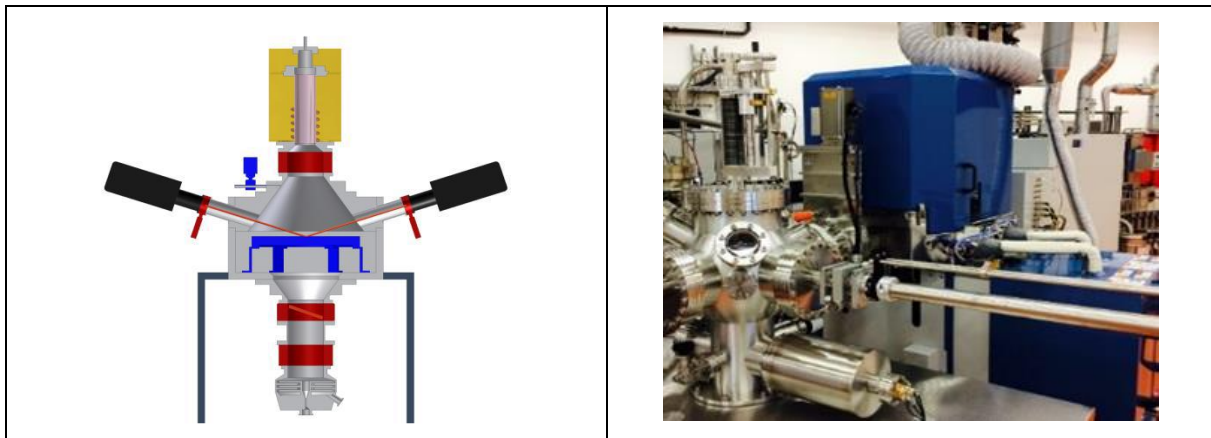


Figure 7: Schematic and photograph of the FlexAl<sup>TM</sup> remote plasma system at HNF cluster

## (b) Photo-Lithography

Photolithography or optical lithography is a process used in semiconductor device fabrication to transfer a pattern from a photo-mask (also called reticle) to the surface of a substrate. Here, photolithography is just briefly described. Details can be found in various references, e.g. [2].

Photolithography involves a sequence of:

1. substrate preparation
2. photo-resist application



3. soft-baking
4. exposure
5. reversal baking
6. flood exposure
7. developing

#### *Substrate preparation*

Prior the application of photo-resist all residuals on the surface need to be removed by appropriate cleaning steps dependent on the substrate material. Special attention needs to be paid to moisture. Therefore silicon wafers are heated and coated with HDMS (hexamethyldisilazane) which promotes a better adhesion between the photo-resist and the wafer.

#### *Photo-resist application*

The wafers are usually covered with photo-resist by spin coating. Depending on the type of the resist and the functionality it should serve (dry etching, wet etching, long or short etching time etc.) the film thickness varies between 0.5  $\mu\text{m}$  up to several micrometer. Typically AZ 5214E photo-resist is spin coated at 4000 rpm to reach a film thickness of 1.4  $\mu\text{m}$  followed by a soft bake at 90 °C for 5 minutes to drive off excess solvents.

#### *Exposure*

In general photoresists are classified into two groups, positive resists and negative resists.

- A positive resist is a type of photo-resist in which the portion of the photo-resist that is exposed to light become soluble to the photo-resist developer and the portion of the photo-resist that is unexposed remains insoluble to the photo-resist developer.
- A negative resist is a type of photo-resist in which the portion of the photo-resist that is exposed to light becomes relatively insoluble to the photo-resist developer. The unexposed portion of the photo-resist is dissolved by the photo-resist developer.

Within the exposure the wafer with photo-resist on top and the photo-mask are brought into contact by use of a so called “aligner”. The photo-mask is made of quartz glass plate on which the structures are patterned in the chromium layer. The structures need to be transferred to the wafer. This assemble is than irradiated with UV light, whereby in the non-covered parts the resist is exposed to the UV light. Note that for the cross-point structures the top electrode pattern has to be aligned to the bottom electrode pattern.

#### *Image reversal*

AZ® 5214 E is a special photo-resist capable of image reversal. Exposed areas may be selectively cross-linked by a reversal bake at 120°C for 45 sec after the first exposure. A flood exposure of the full wafer converts formerly unexposed areas to soluble material, resulting in a negative tone image. The profile generated by such processing has a negative shape (wall profile) allowing for use in lift-off techniques. The schematic diagram of lift-off process is shown in Figure 8.

### Developing

To release the soluble structures, the wafers are rinsed with in a developer. Developers often contain sodium hydroxide (NaOH). Also metal-ion-free developers such as tetramethylammonium hydroxide (TMAH) are used. By the use of a positive photoresist the developer removes all exposed areas, if a negative resist is used all areas which were unexposed in the first exposure step are removed.

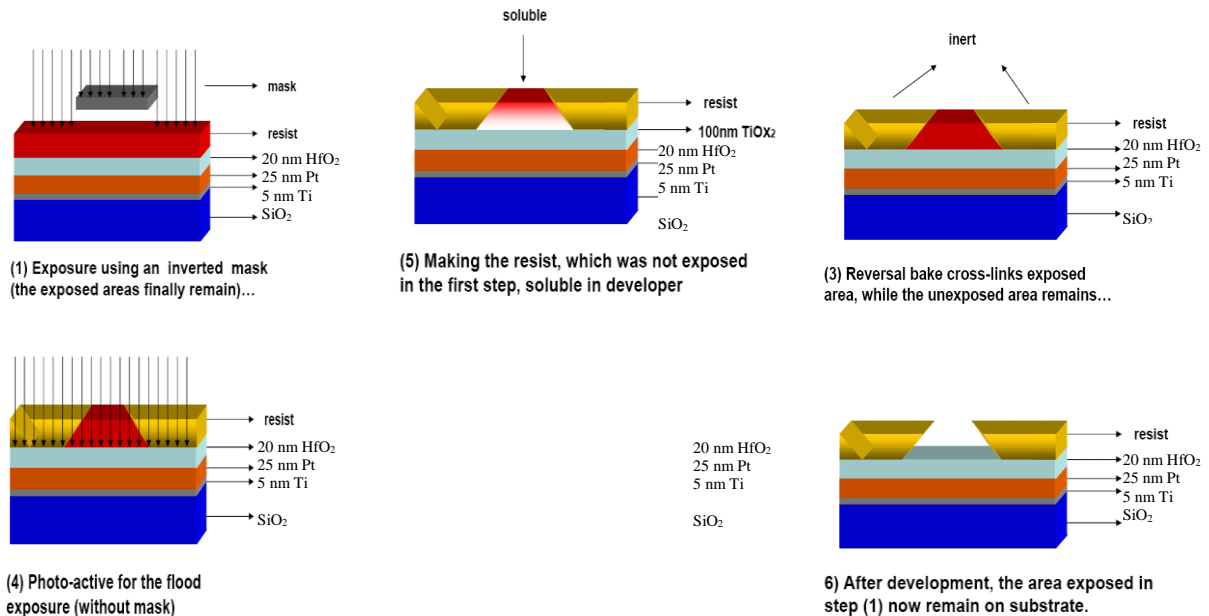


Figure 8: Photolithography - sequence of image reversal processing.

### Lift-off process

For preparation of the top electrode a Ti/Pt metal film will be sputtered over the whole chip area and patterned using the "Lift-off" process (Figure 9.1). During the lifting-off, the photo-resist under the metal layer is removed with the solvent, taking the metal film on top with it, and leaving behind only the metal structures which were deposited directly on the HfO<sub>2</sub> layer (Figure 9.2).

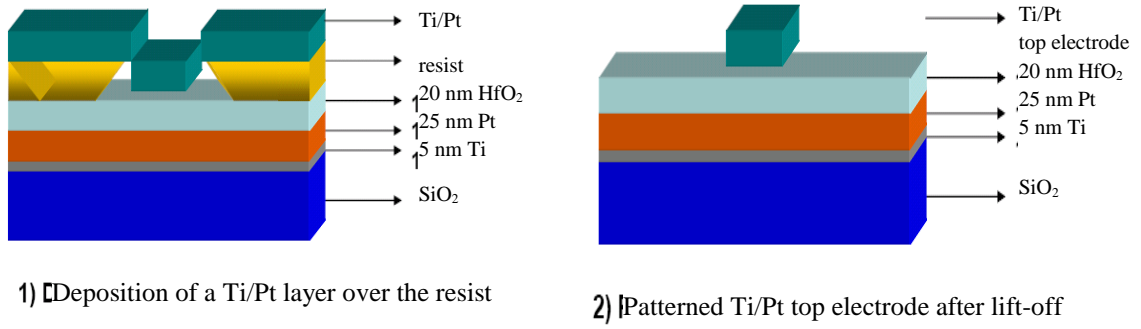


Figure 9: Schematic of “lift-off” technique.

### (c) Sputter Deposition of Metal Films

Sputtering belongs to the physical vapor deposition (PVD) techniques, as the thin film is deposited by condensation out of the vapor phase of the desired material. The sputtering system that will be used for the metal layer deposition is a Univex 450C from Leybold. This sputter machine is designed as a cluster tool with a load-lock and several chambers equipped with different materials like Pt or Ti, each connected to the transfer chamber. The setup enables a consecutive in situ deposition of different materials. For the fabrication of Pt and Ti thin films a DC-magnetron source will be used as depicted in Figure 10.

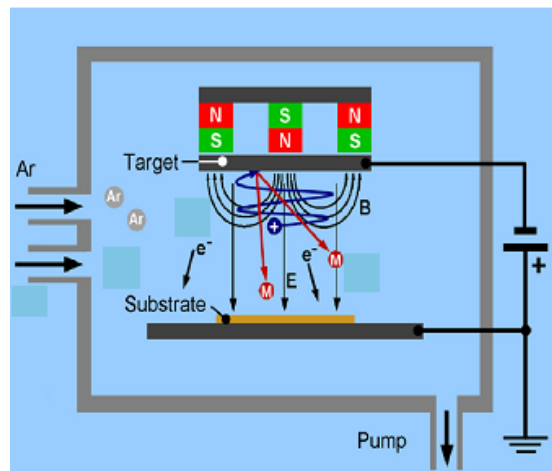
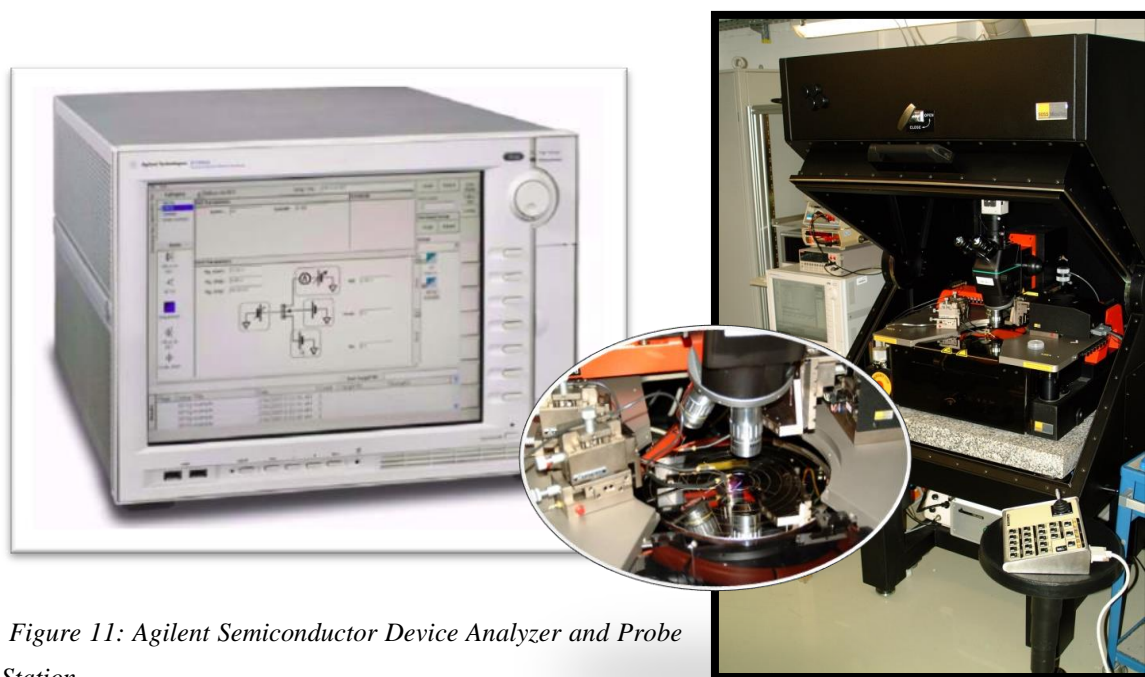


Figure 10: Schematic drawing of the DC- magnetron sputtering process. [5]

This is adequate for the pure metal targets, which are connected to the negative potential of the generator serving as cathode. The grounded substrate support establishes a high electrical field with the target. Pure Ar is injected into the chamber resulting in a working pressure between  $5.1$  and  $13.5 \cdot 10^{-3}$  mbar. This pressure and the high electric field ignite an  $\text{Ar}^+$  plasma, which is also termed glow discharge, due to its relatively low ionization degree. Under the given parameters and a DC input power of  $375$  W a deposition rate of  $2.5 \text{ \AA} \cdot \text{s}^{-1}$  for Pt can be achieved. A stack of  $5\text{ nm}$  Ti and  $25 \text{ nm}$  Pt will be deposited onto the sample with the structured resist pattern to obtain the final cross point electrode structure after accomplishing the ‘lift-off’ process (see Fig. 9).

#### (d) Electrical Characterization of $\text{HfO}_2$ based ReRAM Devices

Besides the surface- and structural analysis with SEM, XRD, AFM etc. (not part of this course, more information can be found in [3]), the electrical characterization is one major part to investigate the material properties in order to realize materials for the resistive RAM. The typical test sample is shown in Figure 4, where the functional film (e.g.  $\text{HfO}_2$ ,  $\text{Ta}_2\text{O}_5$ ,  $\text{TiO}_2$ ) is sandwiched between the lines of bottom and top electrodes of different width.

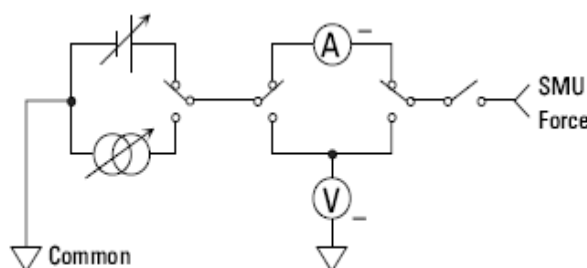


*Figure 11: Agilent Semiconductor Device Analyzer and Probe Station*

The active area during the electrical measurements is thereby reduced to the area of the cross junction. Since the lateral size of the contact point is a factor of  $100$  to  $1000$  larger than the film thickness (typically  $10\text{--}20 \text{ nm}$ ) one may neglect fringe effects. The electrical testing is performed using a Semiconductor Device Analyzer in combination with a Probe Station in order to achieve good connection and sufficient shielding against interferences from the environment (Fig. 11). The Agilent B1500A provides the DC voltage/current output capability,

the DC voltage/current measurement capability, and the AC signal output and impedance measurement capability. So you can perform the current-voltage sweep measurement and the capacitance voltage sweep measurement for example by one instrument. Since the novel memory type relies on the fact that each cell can have two different resistance states, indicating the digital “0” and “1”, and the change between those different states is performed by exceeding a threshold voltage, the most important measurements are I(V)-characteristics. The analyser provides four so called source/monitor units (SMU), which can apply DC voltage or current, and can measure DC current or voltage.

Figure 12 shows a simplified SMU circuit diagram. The SMU can perform the following operations: The SMU has a compliance feature that limits output voltage or current to prevent damage to the device under the test. When the SMU applies voltage, you can specify the current compliance. When the SMU applies current, you can specify the voltage compliance.



- Apply voltage and measure current
- Apply current and measure voltage
- Apply voltage and measure voltage
- Apply current and measure current

Figure 12: Simplified SMU Circuit Diagram [5].

The characteristic parameters of a typical current vs. voltage I(V) measurement are shown in Figure 13. A measurement begins with setting the electrical potential to a start value and holding it for a certain time in order to decay relaxation effects and to get stable measurement condition. Then the voltage is increased stepwise and after a delay time ( $t_{\text{delay}}$ ) the current is measured during the integration time ( $t_{\text{integrate}}$ ). This procedure is repeated until the stop value is reached (single sweep) or the stop value is reached and the voltage is decreased again until the starting voltage (double sweep). All parameters like hold time, delay time, start and stop potential, step height, number of steps etc. are free programmable and depend strongly on the sample material and type of conducting mechanism.

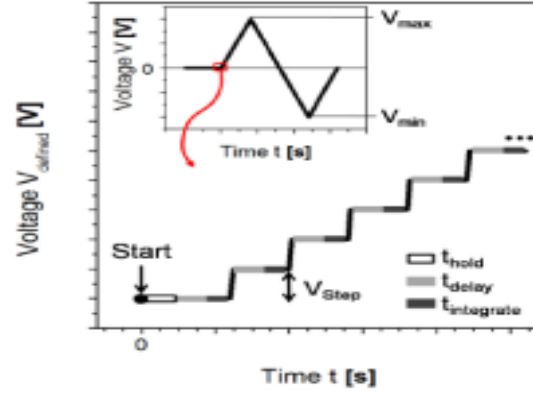


Figure 13: Characteristics parameters of a voltage sweep measurement [5], [6].

Some typical measurement results are presented in Figure 14(a), (b). Initially, the sample is in “off” state corresponding to a low current (Fig. 14a) and high resistance (Fig. 14b). By exceeding a threshold voltage of +0.5 V the resistance decreases dramatically and the current is limited to the current compliance of 100  $\mu$ A.

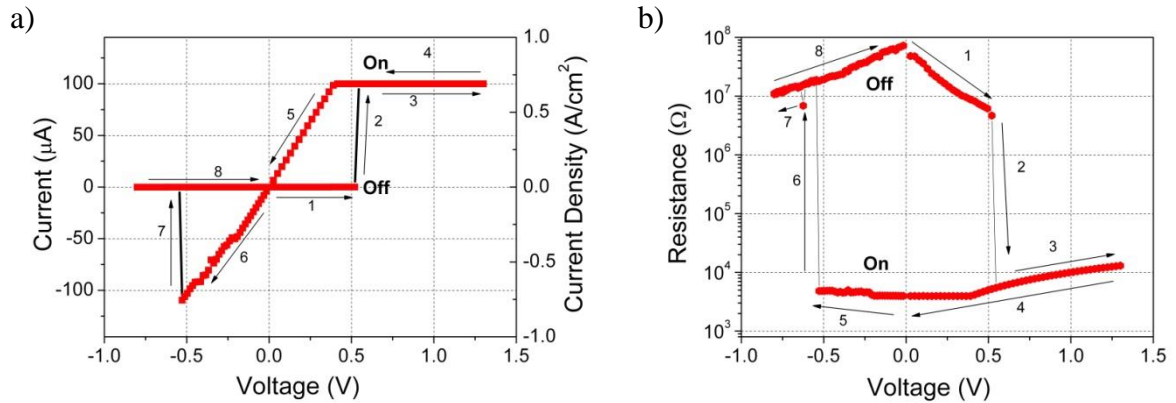


Figure 14: a) Current–Voltage characteristic:  $I(V)$  and b) Resistance–Voltage characteristic:  $R(V)$  of the  $\text{HfO}_2$  based ReRAM device.

Decreasing the voltage results in a hysteretic behaviour (arrow 5) indicating a higher current and a low resistance. Reversing the potential leads to an abrupt decrease of the current when exceeding -0.5 V, which means that the sample switches back from the low resistance state (LRS) to the high resistance state (HRS) corresponding to the change from “1” to “0”. Since the different states are stable without any applied voltage this memory technology is non-volatile.

## 6) Conclusion

In conclusion, the ReRAM devices show the potential to succeed as the next generation non-volatile memory devices. These are easy to fabricate, consume less power and fast in write/read

operation. In addition, the possibility of CMOS back-end integration makes this technology very attractive. However, further efforts are required to realize them in products.

## References

- [1] “Atomic Layer Deposition” by M. Ritala, M. Leskelä, in Handbook of Thin Film Materials - Vol. 1, ed. H.S. Nalwa, Academic Press, (2002).
- [2] Lithography - Theory and Application of Photo-resist, Developers, Solvents and Etchants, MicroChemicals (2006).
- [3] R. Waser, Nanoelectronics and Information Technology, Wiley - VCH, third edition (2012).
- [4] R. Waser et al. “Redox-Based Resistive Switching Memories- Nonionic Mechanisms, Prospects, and Challenges in Advanced Materials, 21, 2623 (2009)
- [5] Operation manual Agilent Semiconductor Device Analyzer B1500A.
- [6] Christian Nauenheim, PhD Thesis (online available), RWTH Aachen University, (2009).