Turun yliopisto
University of Turku

Fabrication and Pseudo-Analog Characteristics of Ta_2O_5-Based ReRAM Cell

Mika Grönroos

2016

Master's thesis

University of Turku

Department of Information Technology

Supervisors:

Prof. Takahashi Yasuo

Prof. Arita Masashi

Ph.D. Lehtonen Eero

Examiners:

D.Sc. Laiho Mika

Ph.D. Punkkinen Risto

TURUN YLIOPISTO
Informaatioteknologian laitos

Mika Grönroos
Ta_2O_5-perustuvan ReRAM-muistisolun valmistus ja pseudoanaloginen toiminta

Diplomityö, 99 sivua
Mikroelektroniikka
Heinäkuu 2016

Memristori on yksi elektroniikan peruskomponenteista vastuksen, kondensaattorin ja kelan lisäksi. Se on passiivinen komponentti, jonka teorian kehitti Leon Chua vuonna 1971. Kesti kuitenkin yli kolmekymmentä vuotta ennen kuin teoria pystyttiin yhdistämään kokeellisiin tuloksiin. Vuonna 2008 Hewlett Packard julkaisi artikkelin, jossa he väittivät valmistaneensa ensimmäisen toimivan memristorin.

Memristori eli muistivastus on resistiivinen komponentti, jonka vastusarvoa pystytään muuttamaan. Nimensä mukaisesti memristori kykenee myös säilyttämään vastusarvonsa ilman jatkuvaa virtaa ja jännitettä. Tyypillisesti memristorilla on vähintään kaksi vastusarvoa, joista kumpikin pystytään valitsemaan syöttämällä komponentille jännitettä tai virtaa. Tämän vuoksi memristoreita kutsutaankin usein resistiivisiksi kytkimiksi.

Resistiivisiä kytkimiä tutkitaan nykyään paljon erityisesti niiden mahdollistaman muistiteknologian takia. Resistiivisistä kytkimistä rakennettua muistia kutsutaan ReRAM-muistiksi (lyhenne sanoista *resistive random access memory*). ReRAM-muisti on Flash-muistin tapaan haihtumaton muisti, jota voidaan sähköisesti ohjelmoida tai tyhjentää. Flash-muistia käytetään tällä hetkellä esimerkiksi muistitikuissa. ReRAM-muisti mahdollistaa kuitenkin nopeamman ja vähävirtaiseman toiminnan Flashiin verrattuna, joten se on tulevaisuudessa varteenotettava kilpailija markkinoilla.

ReRAM-muisti mahdollistaa myös useammin bitin tallentamisen yhteen muistisoluun binäärisen (”0” tai ”1”) toiminnan sijaan. Tyypillisesti ReRAM-muistisolulla on kaksi rajoittavaa vastusarvoa, mutta näiden kahden tilan välille pystytään mahdollisesti ohjelmoimaan useampia tiloja. Muistisoluja voidaan kutsua analogisiksi, jos tilojen määrää ei ole rajoitettu. Analogisilla muistisoluilla olisi mahdollista rakentaa tehokkaasti esimerkiksi neuroverkkoja. Neuroverkoilla pyritään mallintamaan aivojen toimintaa ja suorittamaan tehtäviä, jotka ovat tyypillisesti vaikeita perinteisille tietokoneohjelmille. Neuroverkkoja käytetään esimerkiksi puheentunnistuksessa tai tekoälytoteutuksissa.

Tässä diplomityössä tarkastellaan Ta_2O_5-perustuvan ReRAM-muistisolun analogista toimintaa pitäen mielessä soveltuvuus neuroverkkoihin. ReRAM-muistisolun valmistus ja mittaustulokset käydään läpi. Muistisolun toiminta on harvoin täysin analogista, koska kahden rajoittavan vastusarvon välillä on usein rajattu määrä tiloja. Tämän vuoksi toimintaa kutsutaan pseudoanalogiseksi. Mittaustulokset osoittavat, että yksittäinen ReRAM-muistisolu kykenee binääriseen toimintaan hyvin. Joiltain osin yksittäinen solu kykenee tallentamaan useampia tiloja, mutta vastusarvoissa on peräkkäisten ohjelmointisyklien välillä suurta vaihtelevuutta, joka hankaloittaa tulkintaa. Valmistettu ReRAM-muistisolu ei sellaisenaan kykene toimimaan pseudoanalogisena muistina, vaan se vaati rinnalleen virtaa rajoittavan komponentin. Myös valmistusprosessin kehittäminen vähentäisi yksittäisen solun toiminnassa esiintyvää varianssia, jolloin sen toiminta muistuttaisi enemmän pseudoanalogista muistia.

Avainsanat: memristori, ReRAM, analoginen muisti, neuroverkot

UNIVERSITY OF TURKU
Department of Information Technology

Mika Grönroos
Fabrication and Pseudo-Analog Characteristics of Ta_2O_5-Based ReRAM Cell

Master's thesis, 99 pages
Microelectronics
July 2016

The memristor is one of the fundamental circuit elements in addition to a resistor, capacitor and an inductor. It is a passive component whose theory was postulated by Leon Chua in 1971. It took over 30 years before any known physical examples were discovered. In 2008 Hewlett Packard published an article where they manufactured a device which they claimed to be the first memristor found.

The memristor, which is a concatenation of memory resistor, is a resistive component that has an ability to change its resistance. It can also remember its resistance value without continuous current or voltage. Typically, a memristor has at least two resistance states that can be altered. This is the reason why memristors are also called resistive switches.

Resistive switches can be used in memory technologies. A memory array that has been built using resistive switches is called ReRAM (resistive random access memory). ReRAM, like Flash memory, is a non-volatile memory that can be programmed or erased electrically. Flash memories are currently used e.g. in memory sticks. However, compared to Flash, ReRAM has faster operating speed and lower power consumption, for instance. It could potentially replace current memory standards in future.

A ReRAM memory cell can also store multiple bits instead of binary operation ("0" or "1"). Typically there exists multiple intermediate resistance states between ReRAM's limiting resistances that could be utilized. Such memory could be called analog, if the amount of intermediate states is not limited to discrete levels. Analog memories make it possible to build artificial neural networks (ANN) efficiently, for instance. ANNs try to model the behaviour of brain and to perform tasks that are difficult for traditional computer programs such as speech recognition or artificial intelligence.

This thesis studies the analog behaviour of Ta_2O_5-based ReRAM cell. Manufacturing process and measurement results are presented. The operation of ReRAM cell is rarely fully analog as there exists limited amount of intermediate resistance states. This is the reason why operation is called pseudo-analog. Measurement results show that a single ReRAM cell is suitable for binary operation. In some cases, a single cell can store multiple resistance values but there exists significant variance in resistance states between subsequent programming cycles. The proposed ReRAM cell cannot operate as pseudo-analog ReRAM cell in itself as it needs an external current limiting component. Improving the manufacturing process should reduce the variability such that the operation would be more like a pseudo-analog memory.

Keywords: memristor, ReRAM, analog memory operation, artificial neural network

トゥルク大学
大学院情報科学研究科

Mika Grönroos:
Ta_2O_5抵抗変化型メモリの作製とアナログメモリ動作

修士論文, 99 頁
マイクロエレクトロニクス
2016 年 7 月

メモリスタは、抵抗、容量、インダクタの3回路要素に加えて、もう一つの回路要素として話題になった、Leon Chuaが1971年に提案した受動素子である。30年以上もこの素子は見つけられていなかったが、2008年にHewlett Packard社が、最初のメモリスタ素子を発見したという論文を出した。

メモリスタはメモリとレジスタを合体したような語であるが、その抵抗が可変な素子である。また、その抵抗は、単調な電圧と電流の関係で示されない特徴がある。典型的なメモリスタ素子は、少なくとも 2 つの抵抗値を取り、その間で変化するので、レジスティブスイッチ（抵抗スイッチ）とも呼ばれている。

レジスティブスイッチは、メモリとして利用で、アレイ状に配置したものは、ReRAM (resistive random access memory)と呼ばれている。ReRAMは、メモリスティックなどに使われているFlashメモリと同様の機能を持ち、不揮発性（電源のサポート無しにメモリ状態を保持できる）で、電気的に書き換えが可能である。加えて、ReRAMは、Flashと比べて、高速動作と低消費電力の特徴を併せ持つ。したがって、将来のメモリスタンダードを書き換える可能性を秘めている。

ReRAMセルは、従来の 2 値のバイナリ情報(”0” or ”1”)だけではなく、多値の情報を記憶することも可能である。すなわち、ReRAMの持つ抵抗制限の範囲で、いくつかの中間的な抵抗値を取ることができ、この中間的抵抗値が連続に選択可能になれば、アナログメモリと呼ぶことが可能となる。このようなアナログメモリは、効果的な人工ニューラルネットワークを構成することを可能にする。人工ニューラルネットワークは、人間の脳のように、従来のコンピュータが苦手とする言葉や顔の認識などを可能にする。

本論文は、Ta_2O_5 を用いたReRAMのアナログ特性について研究したものであり、その作製方法と評価結果を示してある。ReRAMの動作は、フルにアナログにはならずに、限られた中間状態を取る准アナログと呼ばれる状況にある。実際の測定結果は、ReRAMがバイナリ動作に向いていることを示した。いくつかの場合、 1 個のReRAMセルが多値を取る結果が得られているが、連続する繰返しの中でも、抵抗値がばらつく現象が見られる。試作したReRAMセルは、准アナログ素子としては動作しなかったが、電流制限素子の付加で改善できる可能性がある。更なる作製プロセスの改善で抵抗ばらつきを減らすことで、准アナログ的な動作が可能になると考えられる。

キーワード: メモリスタ、 ReRAM、アナログメモリ動作、ニューラルネットワーク

Table of Contents

List of Symbols and Acronyms

The following lists the symbols and the acronyms used in this thesis. If some symbol or acronym has other meaning or is missing from this listing, its definition should become clear from the context.

$\rightarrow$	Material implication operation in Boolean algebra
$\wedge$	Logical AND operation in Boolean algebra
$\vee$	Logical OR operation in Boolean algebra
$\neg$	Logical NOT operation in Boolean algebra
a_i	Input parameter of ANN
A_{mem}	Area of memory matrix
α	Shape parameter of parabel at Fermi level
b_j	Output parameter of ANN
B	Subset {0,1} of real numbers
$\boldsymbol{B}$	Magnitude of magnetic field
C	Capacitance
e	Elementary charge
e^-	Electron
E	Electrical potential energy
E_b	Bandgap energy
E_c	Energy of conduction band
E_f	Fermi energy
E_t	Trap level energy
E_v	Energy of valence band
ϵ_0	Vacuum permittivity
ϵ_r	Relative permittivity
F	Fabrication feature size

G	Conductance
G_0	Quantized unit of electrical conductance
h	Planck's constant
i	Electric current
J	Electric current density
L	Inductance
M	Memristance
μ	Electron mobility
μ_V	Average mobility of charge carriers
n_0	Concentration of free carriers in thermal equilibrium
N	Number of conducting filaments
ω_{in}	Angular input frequency
φ, φ_B	Magnetic flux or flux-linkage
ϕ	Hole diameter of ReRAM cell
Φ	Potential barrier height with respect to Fermi level
q	Electric charge
R	Resistance
R_{int}	Intermediate resistance level
R_{OFF}	Off-resistance (or memristance)
R_{ON}	On-resistance (or memristance)
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^n$	n-dimensional set of real numbers
$\boldsymbol{S}$	Surface of vector area
σ	Non-linear function in ANN
t	Time
t_{ox}	Oxide thickness of ReRAM cell
θ	Ratio of the free carrier density to total carrier density
v	Electrical voltage
v_G	MOSFET gate voltage
v_D	MOSFET drain voltage
v_S	MOSFET source voltage
W	Memductance
$w_{i,j}$	Synaptic weight (or strength) in ANN

w	Memristor's state variable
$\boldsymbol{w}$	n-dimensional state variable
$\dot{\boldsymbol{w}}$	n-state derivative of state variable

ALD	Atomic layer deposition
ANN	Artificial neural network
CAD	Computer aided design
CBRAM	Conductive bridge random access memory
CMOL	A hybrid CMOS and resistive switch circuit
CMOS	Complementary metal-oxide semiconductor
CMP	Chemical mechanical planarization
CMR	Colossal magnetoresistive (device)
CNN	Cellular neural network
DC	Direct current
DRAM	Dynamic random access memory
ECM	Electrochemical metallization effect
FPGA	Field programmable gate array
HDD	Hard disk drive
HMDS	Hexamethyldisilazane
HRS	High resistance state
ICP	Inductively coupled plasma
LRS	Low resistance state
MLC	Multi-level cell
MOSFET	Metal–oxide–semiconductor field-effect transistor
MRAM	Magnetic random access memory
NVM	Non-volatile memory
OxRRAM	Transition metal-oxide resistive random access memory
PCM	Phase change memory
PCRAM	Phase change random acceess memory
PLD	Pulse layer deposition
PMC	Programmable metallization cell
PMGI	Polymethylglutarimide
QPC	Quantum point contact

RF	Radio frequency
RIE	Reactive ion etching
RRAM, ReRAM	Resistive random access memory
SCLC	Space charge limited conduction
SEM	Scanning electron microscope
SRAM	Static random access memory
STT-RAM	Spin-torque transfer random access memory
STXM	Scanning transmission x-ray microscope
TCM	Thermochemical memory effect
TEM	Transmission electron microscopy
VCM	Valency change memory effect
WGFMU	Waveform generator/fast measurement unit

1 Introduction

Fundamental physics behind electricity and magnetism are necessary to understand *electronics* and *circuit theory*. Discoveries such as resistance, capacitance and inductance serve as a basis for basic and complex circuit elements. Capacitor – an insulator between two conductors – was invented by German scientist Ewald Georg Von Kleist in 1745. Resistance and resistor were discovered in 1827 by Georg Ohm. Ohm deduced mathematical laws of resistance better known as *Ohm's law* which is one of the earliest concept in circuit theory. Soon after Ohm's findings, inductance, which is used in inductors and transformers was discovered separately in 1831 by two different scientists, Michael Faraday and Joseph Henry.

The three passive fundamental circuit elements: resistor, capacitor and inductor can be defined using four basic circuit variables: *charge*, *current*, *voltage* and *magnetic flux*. Passive elements do not need any external power supply to operate. Active circuit elements such as transistors make it possible to build complex circuits that are currently present everywhere around us but they are not called fundamental circuit elements. However, Leon Chua postulated existence of *the fourth fundamental circuit element* in 1971 [1], that would complete the symmetry of the four basic circuit variables.

Chua named his theoretical component as a *memristor* which is a concatenation of words memory resistor. As the name suggest, a memristor is similar to a non-linear resistor but the fundamental difference is that memristor can change and remember its resistance value even when no external voltage or current is applied to it. This kind of behaviour may be hard to understand intuitively. Also, the strange behaviour in voltage-current plane reveals a pinched hysteresis loop figure. There was no known physical examples of passive memristors for over 30 years after Chua's first publication but in 2008, Hewlett Packard successfully manufactured a device which they claimed to be the first memristor discovered [2]. Soon after their finding many other research groups reported memristor like behaviour in their experimental devices. It has to be noted that memristive

like behaviour was already observed in 1960s and Sharp and Samsung were conducting research on resistive switching memories in early 2000s. However, no one was able or did not try to apply memristive theory to experimental results. Currently memristor theory is a very active research field with many potential applications.

HP's memristor and also discoveries by other research groups are capacitor like thin-film structures where a very thin layer of insulator material is manufactured between metal electrodes. The phenomena that triggers resistance change in materials is called *resistive switching* and many times these components are called *resistive switches* instead of memristors. Currently the exact physical mechanisms that occur inside resistive switches are not known. It is totally possible that there are multiple different mechanisms happening simultaneously which makes it hard to understand this phenomenon.

Naturally the memristor and resistive switches have an application field in memory technologies. Most of the research is towards *resistive random access memory*. Resistive random access memories have many different names in literature: ReRAM, RRAM or CBRAM but essentially they refer to the same thing. In this thesis the term ReRAM is used. Memory technologies can be divided into two categories; *volatile memories* need continuous power to store data and *non-volatile memories* can store their information for a long time after power is cut-off. ReRAM falls into a non-volatile memory category and it could replace current industry standard memory technologies like FLASH-memory that is commonly used e.g. in memory sticks. ReRAM could provide better energy efficiency, faster operating speed and longer retention time in smaller size, for instance [3]. There is also a possibility to store multiple bits in a single ReRAM cell instead of one bit ("0" or "1"). The possible application field of ReRAM is not only limited to memory technologies. There are attempts to use ReRAMs to perform similar computations than transistors perform in logic gates or to provide reconfigurability into analog integrated circuits. One interesting idea is to use ReRAMs for neuromorphic computing purposes like in *artificial neural networks*. Currently software implementations of artificial neural networks are used in image recognition, speech recognition or realizing artificial intelligence, for instance. ReRAMs could provide an efficient way to implement neural networks in hardware level.

Reliable operation is one of the big challenges in realizing resistive switching components. Materials and manufacturing process have an effect on resistive switching which is the reason why no single material combination is superior to any other. ReRAM's voltage-current behaviour as well as voltage impulse response should be analysed in order

to understand its performance suitability for proposed applications.

1.1 Motivation and Approach

The purpose of the research is to achieve pseudo-analog memory characteristics by the use of ReRAMs. Instead of digital operation, pseudo-analog operation allows continuous resistance levels to be utilized in memory. Functionality of such a device is not limited only to memory but it can be also used in computational purposes. The motivation is to realize a ReRAM cell whose resistance levels can be used as connection strength of synapses in artificial neural network systems. The electrical characteristics of one ReRAM cell should be reliable enough to make neural network like structure realistically achievable.

Tantalum pentoxide (Ta_2O_5) based ReRAM cell is proposed to achieve the necessary device requirements for analog memory. ReRAM cell's structure has tantalum top electrode, tantalum pentoxide insulating layer and titanium nitrate bottom electrode. The metal-insulator-metal test structures are fabricated on a silicon wafer with silicon dioxide layer by the use of RF sputtering, lithography and dry etching processes. Different oxide layer thicknesses as well as ReRAM cell areas are studied to see how they affect on ReRAM's electrical properties. The main method is to investigate voltage vs. current characteristics by the use of semiconductor parameter analyser. The measurement equipment provides limitation current that is necessary to prevent ReRAM's breakdown. Limitation current should be changed so that the ReRAM cell is able to achieve hysteresis memory characteristics.

In this study, the first step is to analyse the general characteristics of ReRAM cell. Understanding device's behaviour is crucial in order to realize analog operation. The method to achieve analog operation is to control negative reset voltage by modulating them with voltage sweeps. The next step is to investigate the resistance control by the voltage pulses considering the pulse-coupled neural network. Here the basic control factors are the pulse width and height. If the results are satisfactory, the following step is to control the resistance of ReRAM by the number of pulses, in which the control factor should be appropriately chosen.

Since the semiconductor parameter analyzer's limitation current was found to be insufficient for analog memory, a ReRAM cell with a MOSFET is examined. MOSFET provides another way to control the ReRAM's current by controlling the MOSFET's gate

voltage. MOSFET is also proposed in literature to be a good selector in memory arrays that have a crossbar like structure. Artificial neural networks can be also implemented in crossbar arrays which is the reason that MOSFET-ReRAM structure and its electrical characteristics are studied in the same way as a single ReRAM cell.

1.2 Organization of Thesis

The structure of this thesis is the following:

In Chapter 2 the mathematical background as well as necessary terminology of memristor theory is given. The ideal voltage vs. current characteristics is presented. This chapter is mostly based on Chua's work [1, 4].

Chapter 3 presents the physical processes that occur inside resistive switching. It begins with a short introduction of HP's memristor model and physical explanation. HP's device is chosen because their publication [2], was the first one that connected memristor theory with experimental results. The rest of the chapter presents the terminology and theory of filamentary and interfacial type resistive switching. The switching mechanisms known as redox-based effects that are observed in ReRAMs are also described.

A short introduction of different applications of resistive switching devices is presented in Chapter 4. Digital and analog applications are discussed. The main focus is on resistive random access memories and its comparison to other memory technologies. A very brief introduction to artificial neural networks and their realization using ReRAMs' is also given.

The Chapters 5, 6 and 7 are the practical part of this thesis. First, the manufacturing process is explained in sufficient detail. The general characteristics and measurement results are presented in Chapter 6. The important performance metrics are studied to understand how well the ReRAM cell can operate if analog operation is considered. There is also a brief section about the physical mechanisms that could be present in the ReRAM cell. Finally, the analog operation measurements are presented and analyzed. Chapter 7 describes the manufacturing process of MOSFET-ReRAM structure where measurement results before the final conclusions of this work are given.

In this thesis I manufactured and measured the proposed ReRAM cell. The aim of this thesis is to give a broad overview about memristive theory, physical processes behind resistive switching, ReRAM's practical applications and presenting experimental results of manufactured ReRAM cell. The theoretical background and physical switching

mechanisms are crucial to understand the electrical behaviour of the proposed ReRAM cell. The main scope of this thesis are the measurement results of Ta_2O_5-based ReRAM cells and the analog memory characteristics. If the results are considered good enough for analog memory operation, this work could act as a base for further investigations of ReRAM based artificial neural networks.

2 Fundamental Circuit Elements

2.1 Four Basic Circuit Variables

The elementary charge e is a fundamental physical constant that defines a charge that a single electron (or proton) carries. It is measured in *coulombs* (C) and its approximate value is

$$e = 1.6021766208 \cdot 10^{-19} \text{ C} \tag{2.1}$$

according to NIST Reference on Constants, Units and Uncertainty. Elementary charge must not be confused with electric charge q. Electric charge is a quantity of charge carriers that flow through some cross section in time t

$$q = \int i dt. \tag{2.2}$$

The flow of charge carriers is called an electric current i and it is measured in *amperes* (A)

$$i = \frac{dq}{dt}. \tag{2.3}$$

A quantity of one ampere is also defined by The International System of Units (SI) to be

"*The constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed one meter apart in vacuum, would produce between these conductors a force equal to* $2 \cdot 10^{-7}$ *newtons per metre of length.*"

but current of one ampere can be presented using another SI-unit called coulomb. In that

case, one ampere is one coulomb of charge going through a cross section in one second's time.

The *electric voltage* v is a difference of *electric potential energy* E between two points per unit electric charge. The electric voltage drives charge carriers between two points thus creating a current. In circuits voltage is measured in *volts* (V)

$$v = \frac{E}{q}. \tag{2.4}$$

The fourth basic circuit variable is called *magnetic flux* (or *flux linkage*) φ. The magnetic flux is a time integral of electric voltage

$$\varphi = \int v dt \text{ or } v = \frac{d\varphi}{dt} \tag{2.5}$$

Now the four circuit variables $\{i, q, v, \varphi\}$ have been defined and it can be seen that there exist six different pairwise relations between them

$$\{(v, i), (v, q), (i, \varphi), (\varphi, q), (v, \varphi), (i, q)\}. \tag{2.6}$$

2.2 Three Basic Circuit Elements

Out of six relations presented in (2.6) three of them are well known. The basic *two-terminal passive circuit elements* are *resistor*, *capacitor* and *inductor* whose circuit variable relations are (v, i), (v, q) and (i, φ) respectively. It's not necessary to present any equations of their physical behaviour here, but it is assumed that there exist implicit equations that define their properties. The following functions define the three fundamental circuit elements:

$$\begin{cases} f_R(v, i) = 0 & (2.7) \\ f_C(v, q) = 0 & (2.8) \\ f_L(i, \varphi) = 0 & (2.9) \end{cases}$$

where R, C and L stand for *resistance*, *capacitance* and *inductance* respectively.

The previous three equations generalize *non-linear* fundamental circuit elements but they can be written as voltage or current derivatives of voltage, charge and flux:

$$\begin{cases} dv = R(i)di & (2.10) \\ dq = C(v)dv & (2.11) \\ d\varphi = L(i)di & (2.12) \end{cases}$$

If the values of $R(i)$, $C(v)$ and $L(i)$ are constant, the circuit elements are called *linear*. For example, a *transistor* is not a fundamental element because it is a three terminal externally driven component. On the other hand, a *diode* can be called a fundamental component because it is a non-linear version of resistor thus it belongs to the group of resistors. The unit values for resistance is *ohm* (Ω), for capacitance it is *farad* (F), and for inductance it is *henry* (H).

If (2.2) and (2.5) are written in the form where the current and voltage are integrated over infinite past, relations between inductors (i, φ) and capacitors (v, q) have the following expressions respectively

$$\varphi(t) = \int_{-\infty}^{t} v(\tau)d\tau = \varphi_0 + \int_{0}^{t} v(\tau)d\tau, \tag{2.13}$$

$$q(t) = \int_{-\infty}^{t} i(\tau)d\tau = q_0 + \int_{0}^{t} i(\tau)d\tau, \tag{2.14}$$

where φ_0 and q_0 are initial state of flux and charge at the initial time $t = 0$. The two of these relations, namely those between (v, φ) and (i, q) are not constitutive relations because they are related with (2.13) and (2.14).

Out of six relations only one, namely (φ, q) has not been defined. In Chua's 1971 paper [1], he noticed this missing relation and postulated an existence of the *memristor*.

2.3 The Fourth Element – Memristor

Chua's memristor is a two-terminal passive device which is also known as *flux-charge* memristor. Memristor is the final constitutive relation thus completing the symmetry that is represented in Figure 2.1. The implicit function of memristor is defined in the same manner as in the case of other non-linear circuit elements

$$f_M(\varphi, q) = 0. \tag{2.15}$$

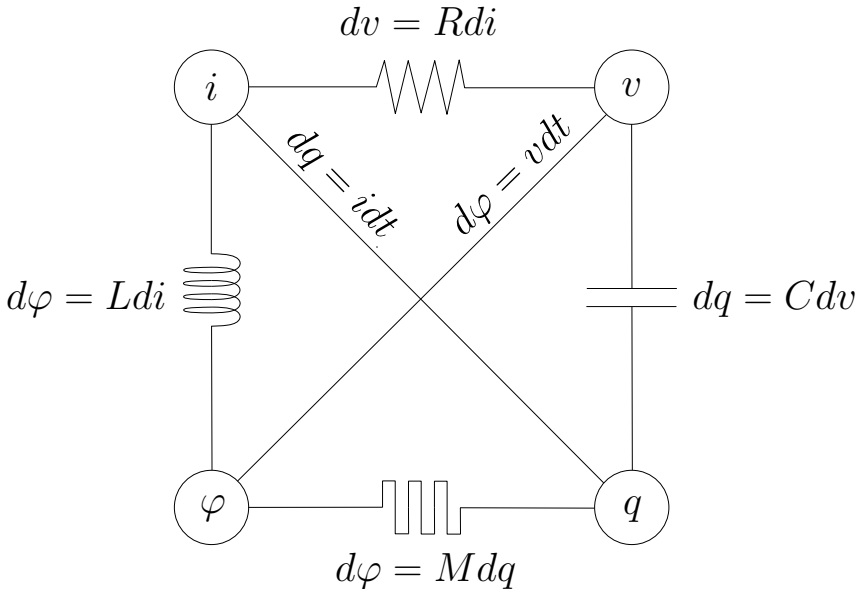

Figure 2.1: Four fundamental circuit elements and their relations to circuit variables. Memristor symbol used in this thesis is shown in the bottom.

The charge derivative presentation is

$$d\varphi = M(q)dq \text{ or } M(q) = \frac{d\varphi}{dq}. \tag{2.16}$$

If (2.16) is substituted with dt, using (2.3) and (2.5) the following relation is obtained

$$M(q(t)) = \frac{d\varphi/dt}{dq/dt} = \frac{v(t)}{i(t)} \tag{2.17}$$

which can be written in the form

$$v(t) = M(q(t))i(t). \tag{2.18}$$

Memristor has a function called *memristance*, M, it's unit value is *ohm* (Ω). The inverse of memristance is called *memductance*, W and it's unit value is *siemens* (S)

$$W(q(t)) = \frac{1}{M(q(t))} = \frac{i(t)}{v(t)}. \tag{2.19}$$

The previous equations are important to understand the $q - \varphi$ and $i - v$ properties of memristor. For example, if M is constant a memristor would be exactly the same as a linear resistor. However non-linear behaviour of M reveals some interesting properties. If (2.5) is used in (2.17) it is easy to see that memristor has a property of remembering

its resistance (memristor is a concatenation of *memory resistor*) which depends on the history of charge that has passed through it

$$v(t) = M\left(\int_{-\infty}^{t} i(\tau)d\tau\right)i(t). \tag{2.20}$$

Memristance keeps changing its value as long as input signal is applied to it. Removing input current (or voltage) makes memristor to keep its state indefinitely. Equation (2.20) defines a *current controlled* memristor.

It is important to understand that Chua's definition of memristor is a classification of *ideal* and *passive* component. Passivity means that component has no internal energy suppliers. Such component must satisfy the following properties in $q-\varphi$ plane [1, 5]:

Criteria 1: *$q-\varphi$ curve must be unique.* Memristor behaves the same way regardless of input waveform. In the same initial condition (q_0, φ_0) memristor changes it's value always the same way if certain amount of charge (or flux) is applied to it.

Criteria 2: *$q-\varphi$ curve must be non-linear.* Non-linearity ensures that memristor differentiates from resistor. If the $q-\varphi$ curve is linear, memristance would be constant and memristor is indistinguishable from resistor. This does not contradict with previously claimed existence of four fundamental elements. A linear memristor belongs to the group of resistors but a non-linear (or linear) resistor does not belong to the group of non-linear memristors thus they are different fundamental elements.

Criteria 3: *$q-\varphi$ curve must be continuously differentiable and increasing.* Continuously differentiable curve guarantees that memristance does not have any discontinuities and it's differentiable in every point thus $M(q), \forall q$ has finite value. Increasing curve $M(q) \geq 0, \forall q$ ensures that ideal memristor will be passive. If the $q-\varphi$ curve is not increasing there is some point q_a, where $M(q_a) = \frac{d\varphi(q_a)}{dq} < 0$, from which follows that voltage can have different polarity than current that is applied to component because instantaneous memristance $M(q(t)) = \frac{v(t)}{i(t)} < 0 \Rightarrow v(t) < 0$ while $i(t) > 0$ (or vice versa). This contradicts with the passivity criteria because such component must have some internal power sources to keep the current flowing through same direction while the voltage over component is inverted. Also instantaneous power dissipated by passive memristor can not be negative, $p(t) = v(t)i(t) = M(q(t))[i(t)]^2 \geq 0$ thus $M(q)$ must be always positive. This criteria could be expanded to monotonically

increasing $M(q) > 0, \forall q$ which limits the $i - v$ curve to have *exactly one* point where $v(t) = 0 \Leftrightarrow i(t) = 0$. Such criteria could be called a *strict passivity* criteria.

Years later after Chua's first theorem of memristor it became clear that a single equation $d\varphi = M(q)dq$ is not enough to explain its physical behaviour in the real world completely. Chua improved his theory on memristors with his former student Kang to generalize memristor to be a subset of *memristive systems* [4]. Memristive systems have input, output and most importantly a n-dimensional *state variable* $\boldsymbol{w}$. Memristor actually needs two equations to characterize it completely. The general equations of memristive systems are

$$y = g(\boldsymbol{w}, u, t)u \tag{2.21}$$

$$\dot{\boldsymbol{w}} = f(\boldsymbol{w}, u, t) \tag{2.22}$$

where functions f and g are continuous, y is the output signal, u is the input signal, t is time and $\dot{\boldsymbol{w}}$ is n-state derivative of $\boldsymbol{w}$. Chua noticed that some previously observed phenomena and devices would classify as memristive system. One example of such device is a component called *thermistor* whose resistance is dependent on its temperature. Another famous example is a circuit model of the North Atlantic squid's giant axon also known as *Hodgkin-Huxley model* [6]. Model's inventors Alan Lloyd Hodgkin and Andrew Huxley were awarded the Nobel Prize in physiology in 1965 and Chua proved that their model had a circuit equivalent where two memristors were used.

As mentioned previously, a passive memristor is a special case in memristive systems. Memristor's state depends on it's previous state and current or voltage input. Because there is no linear relation between current and voltage, memristor should have definition that is related to the physical realization of such device. General nth-order current controlled and voltage controlled memristors are represented by

$$v = R(\boldsymbol{w}, i, t)i \tag{2.23}$$

$$\dot{\boldsymbol{w}} = f(\boldsymbol{w}, i, t) \tag{2.24}$$

and

$$i = G(\boldsymbol{w}, v, t)v \tag{2.25}$$

$$\dot{\boldsymbol{w}} = f(\boldsymbol{w}, v, t) \tag{2.26}$$

Previous implicit equations do not describe behaviour of real physical devices. Most manufactured memristors are thin-film structures and their behaviour can not be categorized the same as flux-charge memristor. Chua mentioned in his memristive systems paper [4], that there are special cases where one-port device is time-invariant.

$$v = R(\boldsymbol{w})i \tag{2.27}$$

$$\dot{\boldsymbol{w}} = f(\boldsymbol{w}, i) \tag{2.28}$$

After HP's discovery in 2008 [2], it became more evident that the previous two equations were very useful in modelling their memristor's behaviour. The detailed equations are presented in Section 3.1 which include process and dimensional parameters. Physical models of two-terminal memristor typically have more limited state variable $w \in \mathbb{R}$.

Previously mentioned criteria and equations guarantee that memristor's $i - v$ curve is a *pinched hysteresis* (Lissajous) curve with a *zero crossing point*. Chua proved mathematically that memristors can not be constructed using other passive fundamental components: resistors, capacitors and inductors.

2.4 Pinched Hysteresis of Memristor

The relations between $q-\varphi$ and $i-v$ are easy to see with help of graphical representations. The continuous pinched hysteresis loop curve is also called the *fingerprint* of memristor. Hysteresis loop gives more intuitive information about memristor's behaviour while it is driven by a sinusoidal input signal. Based on the previously presented criteria memristor's $i - v$ curve has some important properties. It is restricted in 1st and 3rd quadrant and it has only one point, a *zero crossing point*, where both current and voltage are zero.

Figure 2.2 shows an example of memristor's characteristics in $q - \varphi$ and in $i - v$ plane. The $q - \varphi$ curve satisfies all three criteria and it has an initial point (q_0, φ_0). Memristance at point P is a charge derivative of $\varphi(q)$. The pinched hysteresis which is shown in $i - v$ plane is symmetric and has a zero crossing point at the origin. At P is the memristance at

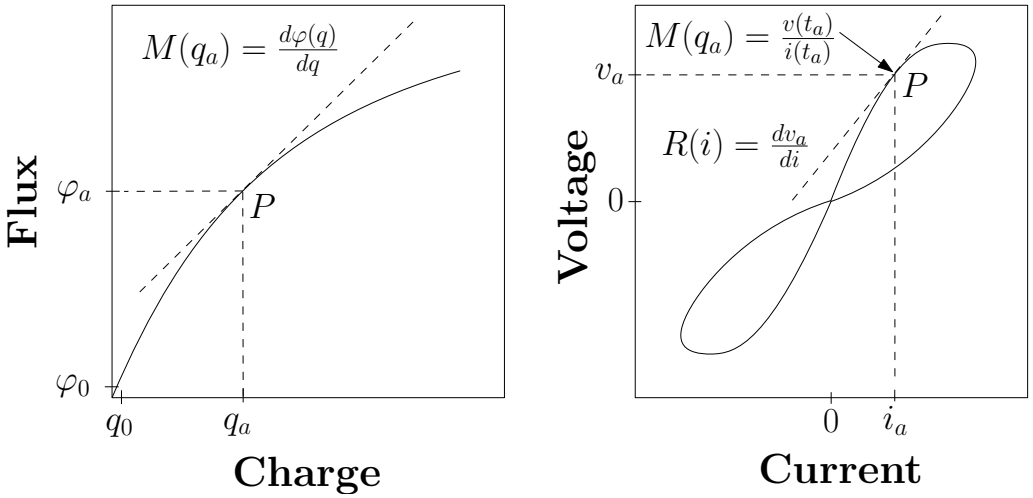

Figure 2.2: $q - \varphi$ and $i - v$ relations of memristor. Adapted from [1].

instantaneous time t_a which is exactly the same as instantaneous resistance. Resistance $R(i) = dv(i)/di$ can be calculated by taking a derivative from the $i - v$ curve at point (i_a, v_a). Memristor's output can not be predicted by looking at it's $i - v$ curve only. Memristor's response is defined by it's $q - \varphi$ curve uniquely and it's state depends on charge that has passed through it.

If one of the criteria is not fulfilled a component would no longer be a memristor. An another example is shown in Figure 2.3. In this case the $q - \varphi$ is unique, non-linear, continuously differentiable but not increasing. At P the memristance $M(q) < 0$ which makes the $i - v$ curve to spread in all four quadrants and there exists three $v = 0$ crossing points. This means that a memristor is no longer passive because such $i - v$ curve is only formed when some internal power source is included.

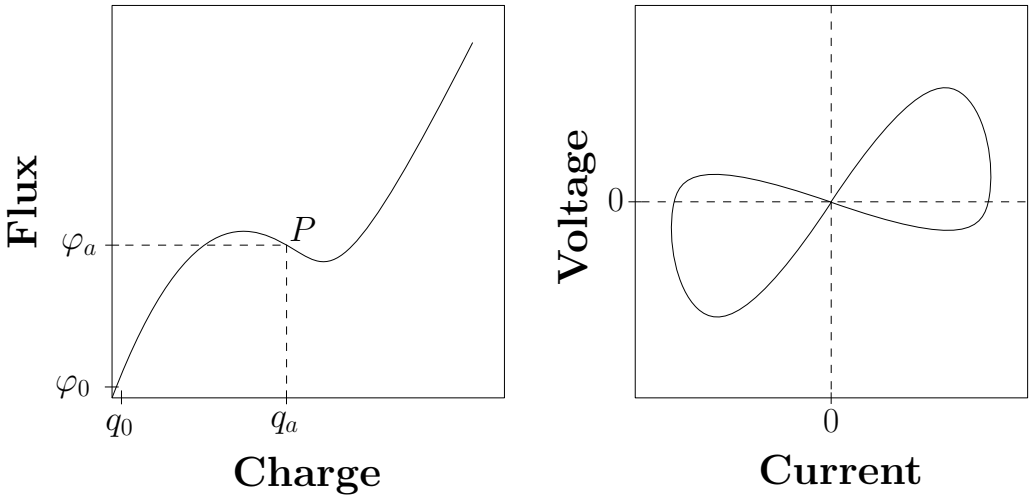

Figure 2.3: $q - \varphi$ and $i - v$ relations of non-memristor. Adapted from [1].

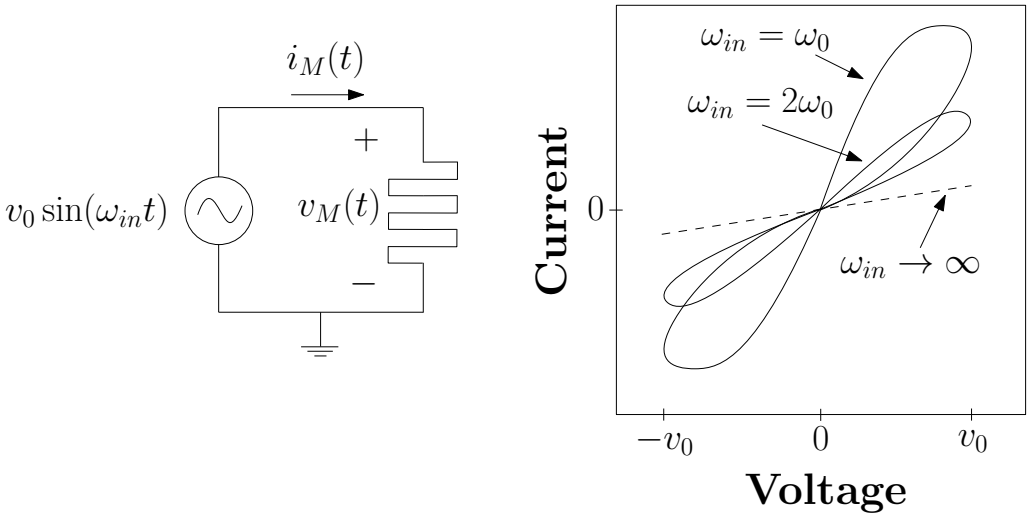

Figure 2.4: Frequency response of memristor. Voltage and current axis are swapped in this figure and subsequent figures to be consistent with existing literature.

The pinched hysteresis loops' waveforms are directly related to input signal frequency. A simple schematic is demonstrated in Figure 2.4 where a memristor is driven by an input $v(t) = v_0 \sin(\omega_{\text{in}} t)$, with amplitude of v_0 and angular frequency ω_{in}. Alternating voltage over memristor at time t is $v_M(t)$ and alternating current through memristor is $i_M(t)$. When input signal's frequency approaches to zero, $\omega_{\text{in}} \to 0$ in other words signal has only dc-value, memristor will eventually saturate to act like a linear resistor. The situation is similar when $\omega_{\text{in}} \to \infty$. The memristor degenerates to a linear resistor but with different resistance value than it is in dc-input's case which is mathematically proved in [4]. Infinitely fast alternating voltage does not allow magnetic flux to change its value because time over one signal period becomes infinitely small. The resistance values depends on a process and materials used on manufactured memristor but basically every *ideal* memristor will have two limiting memristances, R_{ON} and R_{OFF}. The frequency response of memristor shows behaviour where increasing frequency "tightens" hysteresis loops. The $i-v$ curve shows that there are two linear stages between the "transition arcs". These states are typically referred in literature as on-resistance, R_{ON} and off-resistance, R_{OFF}, where $R_{\text{ON}} < R_{\text{OFF}}$. [5]

The mathematical definition of memristor presented in this chapter differs very much from *real* devices. Real memristor do not necessarily have clean pinched hysteresis curves in $i-v$ plane or they are not categorized as true flux-charge memristors. The manufactured devices belong to a totally different subset than memristor in memristive systems. There is ongoing debate in scientific field whether manufactured memristors are real memristors

or are they part of some broad class of resistive systems. One paper published on Nature argues that memristors are impossible to manufacture in real world without using magnetic induction [7], and it may be possible that real memristor will be found in other physical structures than thin-film structures. Chua's mathematical definition is not under criticism but there exists disapproval against HP's device and other thin-film structures because they lack magnetic flux property, yet they are referred as memristors.

Existing literature mixes up terms memristor and memristive systems repeatedly. In this thesis term *memristor* is used although the device does not necessarily meet all the mathematical criteria that is set to it.

3 Physical Realization of Memristor

This chapter gives a brief overview of mechanisms behind resistive switching phenomenon as they are necessary to understand the behaviour and material choices of resistive switches. First, a short introduction of the history behind the switching observations and the HP's memristor model is presented. Further sections should give the basic idea of different switching types and models. Especially the redox-based effects behind resistive switching mechanisms are explained.

There may be a misconception in public that HP's device was the first memristor manufactured. Actually the resistive switching phenomenon was already reported in the 1960s. Some metal oxides such as NiO [8], and Ta_2O_5 [9], were reported to have forming process produced by electrical field and after forming they showed switching properties with two different resistance levels and some memory effects. However, those devices were not robust enough to have any practical use and it took many decades until the interest on resistive switching rouse again among scientific field.

In the early 2000s there were more successful attempts to manufacture resistive switches. Materials like NiO [10], and TiO_2 [11], showed very promising switching properties. In 2002 Sharp published a paper [12], where they presented their *colossal magnetoresistive* (CMR) device that was manufactured using 0.5 µm *complementary metal-oxide-semiconductor* (CMOS) process. Sharp's memory architecture was a 64-bit array where the memory component was connected in series with a transistor or diode. These configurations are called *one-transistor–one-resistor* (1T1R) and *one-diode–one-resistor* (1D1R) correspondingly. The memory cell showed resistive switching when 3 V voltage pulse for programming and erasing was applied. Write and erase times were reported to be 20 ns and 10 ns while peak operating current was 200 µA. Using different pulse widths Sharp's researchers noticed that the device was able to have multiple resistance states which could be used in multi-bit storage. However they did not publish any information of endurance or retention

time. Sharp referred their device as *resistance random access memory* (RRAM) which was the first publication to use this term.

A small breakthrough was also made in 2004 when Samsung published a paper about their resistive memory [13]. Samsung accomplished to manufacture a NiO memory cell on 0.18 µm CMOS process using a 1T1R structure. They referred their memory structure as *transition metal-oxide resistive random access memory* which was shortened to OxRRAM. Samsung made memory technology oriented tests to their device where they showed that voltage pulses were necessary to switch from *low-resistance state* (LRS) to *high-resistance state* (HRS). They used operating voltage below 3 V and controlled current of 2 mA while running 10^6 set/reset cycles and 10^{12} read operations. However Samsung had problems with failing endurance after one million set/reset operation. Also their high temperature baking test resulted many defunct memory cells mainly due soft errors. It is interesting to note that none of the papers about resistive switching did not refer to the original memristor paper. Maybe this was because Chua's papers were full of complex mathematics and his interests did not meet with experimental material scientists.

In 2008 HP Labs finally connected the dots and manufactured a device which they referred as a memristor. Their sandwich like structure used two platinum electrodes between titanium dioxide insulator. It was a beginning of multiple thin-film memristor implementations using different electrode and oxide materials. However it is impossible to give a single comprehensive equation or explanation of how thin-film memristor works because it's switching behaviour is dependant on not only materials used but also their combination [3].

In the next section HP's thin-film memristor is presented because it was the first to connect memristor theory to experimental observations that resulted multiple publications[1] from many research groups that observed resistive switching in metal-oxide nano structures.

3.1 HP's Device Model and Physical Properties

HP Labs presented their physical model of their two terminal memristor-like device in [2]. The first mathematical model of their device is very simple and some improved and more

[1] Search report from *Web of Science* (`http://apps.webofknowledge.com`) results 936 articles between years 1971–2007 and 14 902 articles between years 2008–2015 using keywords "`memristor OR memristive OR resistive switching`".

complex modelling equations have been presented and studied [14,15]. HP's model is presented in Figure 3.1. External sinusoidal voltage $v(t) = v_0 \sin(\omega_{\text{in}} t)$ changes the boundary area between doped and undoped regions. Strukov et al. [2], tried to induce mathematical model of their device which would match Chua's definition for memristive systems. They obtained the following equations

$$v(t) = \left(R_{\text{ON}} \frac{w(t)}{D} + R_{\text{OFF}} \left(1 - \frac{w(t)}{D} \right) \right) i(t) \tag{3.1}$$

$$\frac{dw(t)}{dt} = \mu_V \frac{R_{\text{ON}}}{D} i(t) \tag{3.2}$$

where D is the width of oxide material, μ_V is average ion mobility and formula for $w(t)$ is

$$w(t) = \mu_V \frac{R_{\text{ON}}}{D} q(t). \tag{3.3}$$

Equations (3.1) and (3.2) are modifications of equations $v = R(w)i$ and $\dot{w} = f(w, i)$. Because $R_{\text{ON}} \ll R_{\text{OFF}}$ it follows that inserting $w(t)$ to (3.1) following approximation is obtained

$$v(t) = R_{\text{OFF}} \left(1 - \frac{\mu_V R_{\text{ON}}}{D^2} q(t) \right) i(t). \tag{3.4}$$

These equations show that HP's device belonged to group of resistive systems. They had to show that their device was memristor and it should also have charge dependent memristance. This was indeed the case because memristance can be written in the form $v(t) = M(q(t))i(t)$ which results memristance equation

$$M(q(t)) = R_{\text{OFF}} \left(1 - \frac{\mu_V R_{\text{ON}}}{D^2} q(t) \right). \tag{3.5}$$

It is true that previous equations lack direct connection to magnetic flux that was crucial in defining a memristor. Because there were no clear indications of magnetic fields inside small resistive switching components it was hard for researchers to intentionally manufacture memristive devices. This confusion can be avoided by using a term *flux linkage* instead of magnetic flux as both equations are defined exactly the same in (2.5). Magnetic flux's physical explanation is $\varphi_B = \boldsymbol{B} \cdot \boldsymbol{S}$ where magnitude of magnetic field $\boldsymbol{B}$ passes through surface of vector area $\boldsymbol{S}$. Flux linkage does not represent any physical magnetic field but only a time integral over voltage.

HP's model assumes that R_{OFF}, R_{ON} and D are independent of q and t, which may

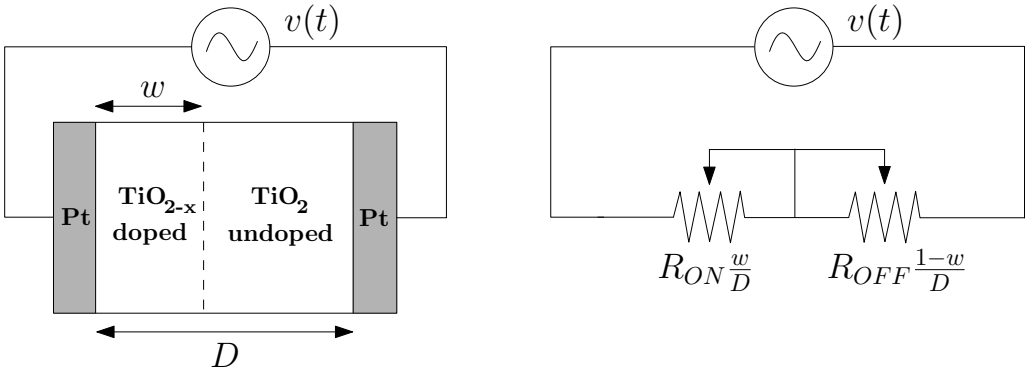

Figure 3.1: A simplified model of HP's memristive device shows a sandwich structure where width of $TiO_{2\text{-x}}$ doped area defines the state variable w. A circuit equivalent is shown on the right where two variable resistances R_{ON} and R_{OFF} represent materials $TiO_{2\text{-x}}$ and TiO_2 respectively.

not be the case in practise. Also HP's claim of memristor discovery led their first model to define memristive behaviour rather than explaining physical mechanisms behind phenomena.

A state variable w in the model has more limited range than Chua's mathematical definition where $\boldsymbol{w} \in \mathbb{R}^n$. Theoretically, a state variable that is a width of doped region could only vary between platinum electrodes in other words $w \in [0, D]$. It is proportional to q that passes through TiO_2 region.

A brief physical explanation of how resistive switching happens in TiO_2 is presented in this section. At the time when HP was doing experiments on their device using micro-scale and nano-scale devices they noticed that memristance was more noticeable in very small structures. Because any *real* component e.g. resistor or capacitor also have parasitics inside them (such as parasitic resistances or parasitic inductances) it may be possible that they contain parasitic memristances also. However this parasitic element is undetectable in large scale devices which made the discovery of memristor to be a very difficult task. As electrical devices started to shrink in size, memristance became more noticeable. There are papers which misinterpreted memristance as tunnel effect [9], or anomalous capacitance [16]. Currently, memristance is known to be a dominant element in nano-scale thin-film metal-oxide structures. In theory, a memristance in nanometer devices is one million times larger than in micrometer devices. This happens because a memristance e.g. in TiO_2 devices has a factor of $1/D^2$ as shown in simplified model in (3.5).

Strukov and Williams investigated switching speeds of $Pt/TiO_2/TiO_{2\text{-x}}/Pt$ device [17]. In their model the motion of atomic particles caused the resistive switching. Charge carriers could be ions, molecules or vacancies. Even relatively small non-stoichiometric of

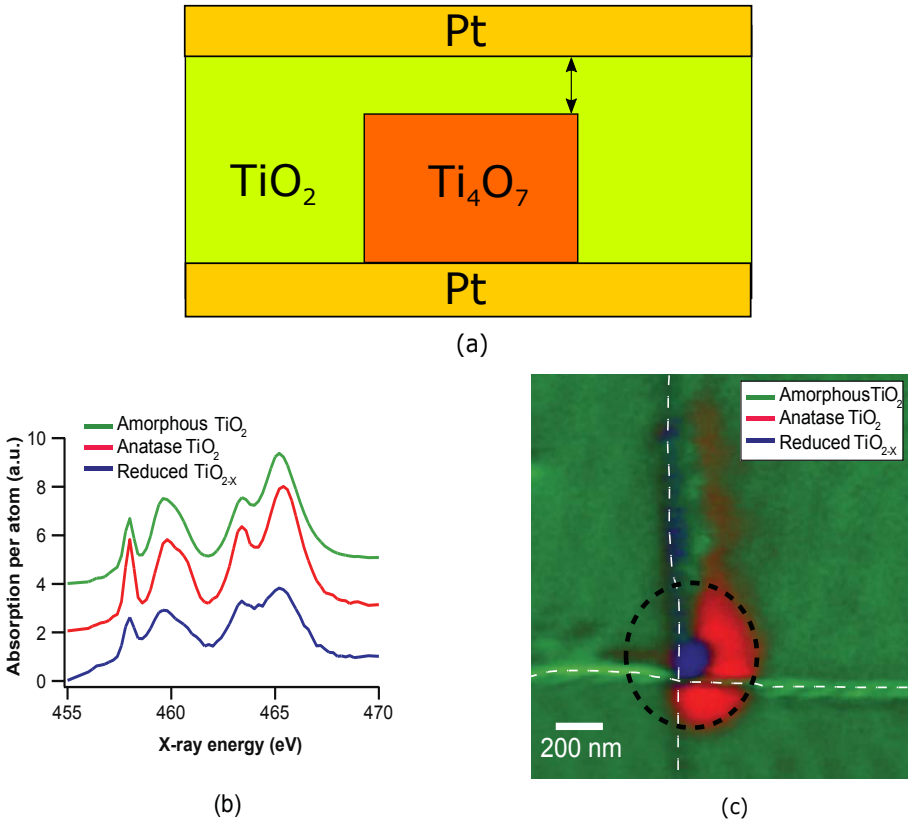

Figure 3.2: (a) A simplified cross cut of the device which shows Ti_4O_7 channel. Resistance switching happens when this channel is modulated using electric field. (b) X-ray absorption spectrum revels three different materials present in device. (c) Colour mapped STXM image of the device shows the formation of conductive channel. Partially taken from [18].

0.1 % in TiO_{2-x} which is equivalent approximately $5 \cdot 10^{19}$ dopants/cm^3 makes the oxygen deficient layer highly conductive [15].

The first manufactured devices used a electroforming step to create a conducting channel in TiO_2 region. Strachan et al. studied [18], forming process and material changes inside the device. Before the forming step the device showed strong rectifying behaviour that could be interpreted as Schottky like contacts at bottom and top interfaces. It turned out that after initial forming step there exists a 10 nm to 200 nm wide channel that would form in a random location inside TiO_2. Researchers probed TiO_2 region with x-ray absorption spectroscopy. The spectrum, which is shown in Figure 3.2, revealed three different phases or sub-oxides of titanium dioxide and the width of conducting channel. Spectral analysis revealed that the material in conducting channel was actually a sub-oxide of TiO_2. *Transmission electron microscopy* (TEM) of reduced TiO_{2-x} region showed that the diffraction pattern figure was the same as the crystal form of Ti_4O_7 which is also known as *Magnéli phase*. The on-off switching happens when the width of region between

Ti_4O_7 and Pt changes. This modulation which is only in range of couple nanometer causes significant changes in resistance in spite of the fact that the conducting channel is never fully shorted with electrode. A *scanning transmission x-ray microscope* (STXM) analysis of the device showed that the formed channel had an approximate width of 100 nm which is color mapped in STXM image in Figure 3.2. The resistive switching mechanism of TiO_2 was confirmed by other independent groups especially Kwon et al. [19]. They concluded that increasing concentration of oxygen vacancies would result strong *Joule heating effect* and high temperatures drive thermodynamic forces to form the Magnéli phase.

3.2 Resistive Switching

When knowledge and observations of resistive switching mechanisms and materials increased, one can argue, that the terms memristance and memristor are good for explaining and defining all the discovered devices. Many times the pinched hysteresis curves are far from being clear and beautiful "butterfly" or "bow-tie" shaped figures. In some devices the pinched hysteresis in two different quarters of $i - v$ plane is totally missing because the switching happens only on positive voltage. However, the fact is that currently new devices and materials are researched extensively and part of this happened because there was an attempt to discover the Chua's fundamental circuit element. Some papers might still refer Chua's 1971 paper but nowadays many authors do not find it necessary as their device and research is more concerned to have use in practical applications instead of proving Chua's theory of postulated memristor. The term *resistive switching* is often used in literature and it is in many cases more illustrative word for the whole phenomenon.

As noted previously, a single comprehensive equation or model that would explain the behaviour of all different thin-film devices is practically impossible. Resistive switching does not only depend on materials that are used but it is also dependent on their different combination [3]. Resistive switching behaviour can be classified into two different types: *unipolar* or *bipolar*.

In unipolar (or non-polar) case, resistive switching happens when amount of applied voltage (or current) exceeds certain level. This means that change from HRS to LRS which is also called as *set* phase and correspondingly the change from LRS to HRS which is called *reset* phase happens in the same voltage polarity. In bipolar switching setting and resetting mechanism does not only depend on amplitude but also polarity of applied

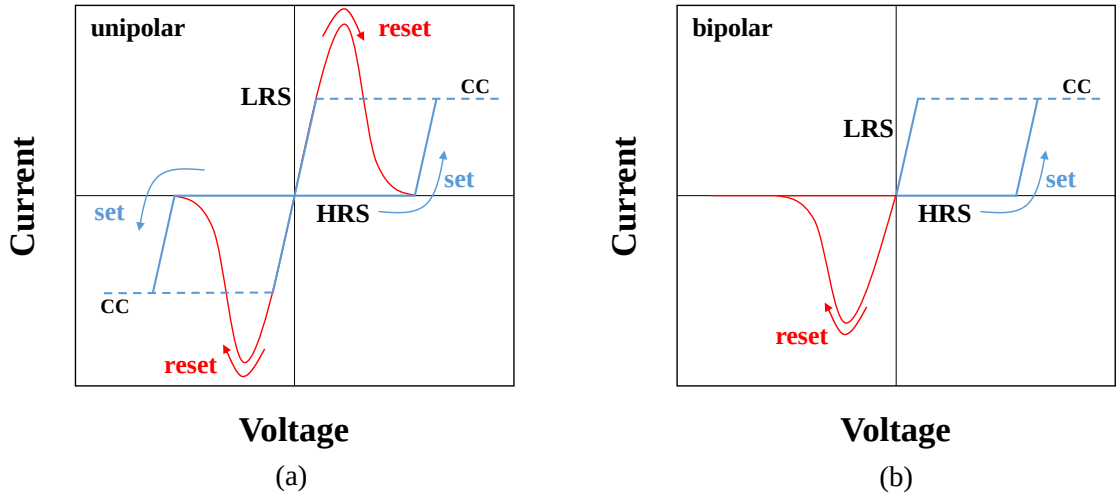

Figure 3.3: The $i - v$ curves of (a) unipolar mode and (b) bipolar mode. In unipolar mode it is possible to change device state between HRS and LRS with both negative or positive voltage. In bipolar mode only the opposite polarity of voltage have the effect of state change. In both cases the current is limited by the compliance current (CC).

voltage. For example, set phase happens on positive voltage when amplitude reaches certain level. Applying increased positive voltage would not cause any change in device's resistance state. However, a subsequent negative voltage will cause a change in device thus resetting it back to its previous state. Figure 3.3 shows the typical $i - v$ curves of two switching modes. Some devices may have the set happening on negative voltage and reset on positive voltage and the curves may be far from being clean hysteresis curves. They are more likely to be distorted or "noisy" between different cycles, so this figure is only used to present and give the understanding of basic behaviour.

Typically a fresh sample of manufactured device has some *initial resistance* which might be different from the resistance states that are present after applying multiple switching cycles [3]. The initial resistance state might need a higher voltage to trigger resistive switching behaviour in device. This is usually called a *forming voltage* and the phenomenon is referred in literature as *forming* or *electroforming*.

The initial forming and following set phases are usually very fast and abrupt processes so it is necessary to prevent device breakdown by limiting the current when device changes it's state to LRS. The *compliance current* is achieved having a transistor, a diode or another resistor in series with resistive switching component. If the manufactured device lacks the mentioned components, compliance current is provided by measurement devices like semiconductor parameter analysers or other corresponding equipment. However, in some cases the switching time is so fast that measurement equipment is not fast enough

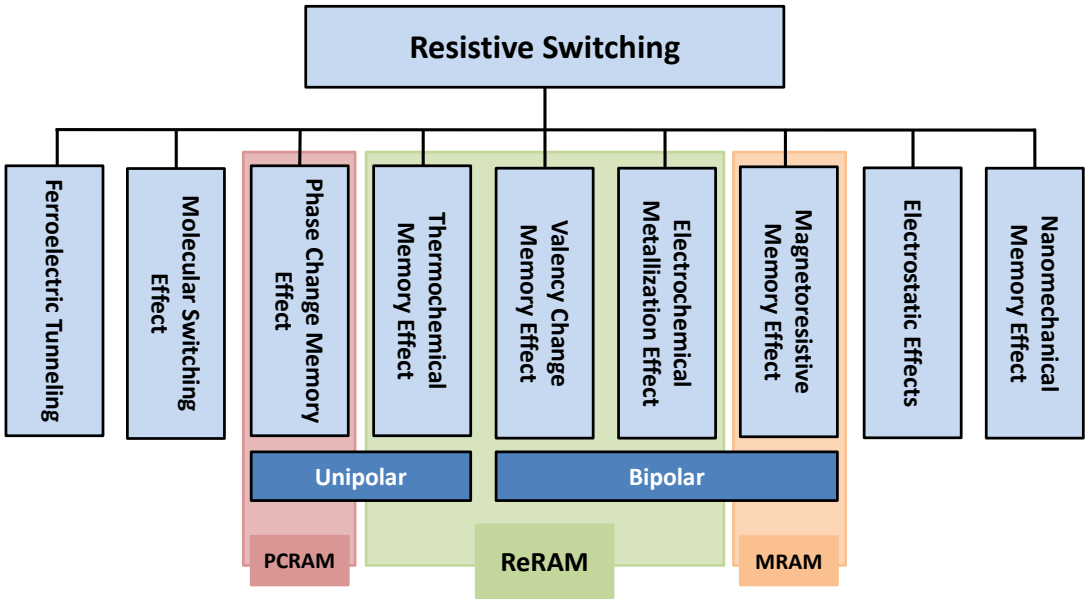

Figure 3.4: Classification of resistive switching mechanisms based on physical driving forces. Unipolar and bipolar operation of resistive random access memories are typically associated with three redox based mechanisms: ECM, VCM and TCM. Adapted from [20].

to react to set process and in such cases the device may break down unintentionally.

Some external ambient conditions such as temperature have also effect in resistive switching. Even the fabrication conditions and process have some effect because the structure of metal-oxide material is never totally ideal and uniform. Possible methods to manufacture oxide based devices include *reactive sputtering*, *pulse laser deposition* (PLD), and *atomic layer deposition* (ALD). Especially the ALD shows promising results because it has an ability to control the uniformity and thickness of the oxide film which reduces the variance between devices as well as variability in single device. [3]

A large variety of different physical phenomena are observed that lead to resistive switching. Waser et al. presented a very comprehensive review [20], where different switching mechanisms are divided into nine categories that are shown in Figure 3.4. Some of the physical driving forces are related to unipolar switching while others contribute to bipolar switching. Waser's paper focused on *redox-related chemical effects* that are typically associated with *resistive random access memories* (ReRAM). A more recent review

about switching mechanisms was introduced by Pan et al. [21], where they focused on ion migration, charge trapping and thermochemical reactions in inorganic and also organic materials. Different switching mechanisms are utilized in different memory technologies such as *phase change random access memory* (PCRAM) or *magnetoresistive random access memory* (MRAM). Redox (reduction-oxidation) based mechanism include *electrochemical metallization effect* (ECM), *valency change memory effect* (VCM) and *thermochemical memory effect* (TCM). These effects are explained in Section 3.4.

3.3 Filamentary and Interface Type Switching Models

In previous section it was mentioned that resistive switching can be classified into two different categories, unipolar and bipolar. The type of conducting path can also be categorized into two types. The proposed models for switching are *filamentary type conducting paths* and *interface type conducting paths*. The general thumb-of-rule link between conduction path and switching behaviour is that the conduction filament can be associated with both types of switching behaviour but interface-type mechanism is typically related only to the bipolar switching behaviour [22]. In Figure 3.5 these two models are shown. However, the details of resistive switching are still missing a comprehensive explanation and it is an active research area. Many different research groups have attempted to see what happens inside the device during switching [18, 19, 20, 23, 24, 25, 26].

3.3.1 Filamentary Type Model

Basically, in filamentary type model the forming process happens when a soft electric breakdown happens in oxide material between cathode and anode. Soft breakdown means that this condition can be reversed even though not fully because in many cases the first forming is reported to require higher voltage than following set forming steps. The filamentary switching process is basically divided into three different sections: forming, set and reset.

Filamentary type switching can be also recognized by it's "fingerprint" of area independence. This means that increasing (or decreasing) device area has almost no effect on the ratio of on-resistance and off-resistance. A simple explanation for this is the formation of single filament during set process. Formation of filament can happen anywhere in ox-

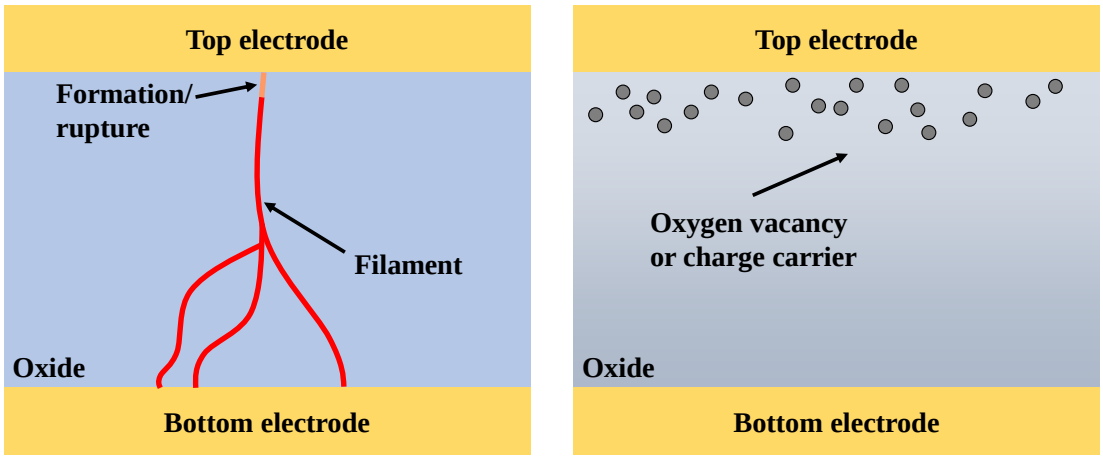

Figure 3.5: Simplified models for (a) filament conduction paths and (b) interface-type conduction paths. Adapted from [22].

ide material and if the filament is expected to be have the same "strength" in consecutive forming, the conducting current is also expected to be the same thus having any noticeable effect on on-resistance value. Increasing device size does not matter if only one filament is allowed to form. In reality, filament's geometry is never identical between consecutive switching cycles. Precise control of filament structure is not a trivial task because of the stochastic nature of filament formation. This is the main reason why $i - v$ curves of manufactured device have high statistical variability between switching cycles.

The oxide thickness is proportional to forming voltage in filamentary type switching. If oxide thickness is reduced strong electric field is not necessary to form a conducting channel in oxide. In most devices the oxide thickness is in scale of nanometers. The dependence between forming voltage and oxide thickness is mostly reported to be linear. Because the initial forming voltage is typically higher that following set voltages, researchers began to question if it is possible to manufacture *forming free* devices. Reports of forming free devices exists e.g. Lee et al. [27], succeeded to manufacture a 3 nm oxide thickness in HfO_x device that did not require initial forming. Fang et al. [28], showed a multilayer HfO_x/TiO_x structure without need of initial forming.

Some models assume that there exists only one conducting filament. This may be possible in some transition metal oxides where the first forming filament is responsible for conducting all electrical current. Still experimental data suggest that there may be hints of multiple filament formations. If multiple filaments are formed the device would also

have multiple resistance states instead of two: on-resistance and off-resistance. This kind of behaviour is beneficial in memory technologies where single device could store multiple bits. Formation of multiple filaments can be achieved using e.g. multiple compliance current levels. Higher current would form several filaments, although rupturing those filaments is usually an uncontrollable process. Repeatability is also an issue when reliable behaviour is necessary but many times the single filament devices suffer from these issues too. The filament breakdown should happen sequentially during voltage sweep in discrete manner. Ideally, the voltage levels where the rupture happens, should be same between sequential sweeps [29]. Multiple resistance states are also possible to achieve with single filament when the size of theinter filament is changed by precise voltage control. In this case the controllable resistance states are hard to achieve because the resistance switching phenomena is highly non-linear. However, if small changes in filament size can be controlled, there is a possibility to achieve pseudo-analog resistance levels.

The conducting filaments in transition metal oxides involve e.g. oxygen vacancy transport. Studying the filament may show defects in local oxygen content as it is shown in [19]. This explains the reason why transition metal oxides are good candidates for resistive switching devices. They have at least two oxidation states and they are good ionic conductors. In unipolar switching the forming process creates a filament consisting of oxygen vacancies. During reset process oxygen atoms fill the vacancies in filament thus breaking it. This model is limited as it does not take into account any electrode materials. The switching needs an active electrode e.g. platinum or copper. In this case, models assume that active electrode catalyses oxidation near the electrode interface. The filament rupture does not happen in the middle of oxide material as the pure oxygen vacancy models indicates. The bipolar switching needs asymmetric structure to operate that is achievable using only one active electrode material (which is also the case in unipolar switching) or using cells with non-uniform built-in oxygen vacancies. The active electrode, typically anode, acts as an oxygen reservoir in forming as oxygen ions drifts towards the electrode thus creating an interfacial layer. In reset, the oxygen ions recombine with the vacancies or oxidize the metal back to high resistance state. This model which was proposed by Yu et al. [30], to be consistent with unipolar and bipolar cases. However, as Wong noted in his review paper [3], the model should be viewed only as a phenomenological description of experimental observation and there exists alternative theories. A more comprehensive review about filament formation and physical processes are given in Section 3.4.

3.3.2 Interface Type Model

Interface type conducting path model is typically used in complex oxide materials such as doped *perovskite* oxides. Perosvkite complex metal oxide means a material that has similar crystal type structure than calcium titanium oxide $CaTiO_3$. As the opposite of filamentary type switching in interface type switching the resistance is expected to be area dependant. This idea can be explained in a simple way because the switching happens in the whole interface that exists between electrode material and oxide material, and thus the resistance depends on the device area. The interfacial switching in perosvikte oxides was studied by Sawa [22]. He made an assumption that the possible origin of resistive switching is the Schottky barrier because the device has a capacitor like structure and the resistance can be changed applying an electric field over it. The switching is an outcome of Schottky barrier modulation. Modulation moves oxygen vacancies towards or away from the interface. Other proposed models that involve e.g. electrochemical migration of oxygen vacancies or trapping of charge carriers are presented in 3.4.

Sawa noticed that some materials would allow resistive switching but changing anode material would lead to non-linear resistor behaviour. This dependence make it more evident that Schottky barrier is indeed the probable cause of contact resistance. For example, a cell made from $Ti/Pr_{0.7}Ca_{0.3}MnO_3$ (shortened to Ti/PCMO) would lead to resistive switching while Au/PCMO showed no switching behaviour. It can be assumed that characteristics of switching depend on electrical properties and some studies would confirm that by modifying the interface properties e.g. using different doping levels of charge carriers. If the amount of charge carriers are reduced by annealing a cell made of Ti/PCMO at 670 K in air or O_2 atmosphere, curves in $i - v$ plane did not show any hysteresis. Finally Sawa noted that the p-type and n-type materials have opposite properties. In n-type material oxygen vacancies act as donors and in p-type material they are acceptors. He concluded that electrochemical redox reaction may indicate that oxygen vacancies are important factor in interface-type switching mechanism. [22].

Interfacial type switching is not only observed in perovskite materials but also in simple binary oxides although the mechanism in latter is supposed to be simpler. For example Yang et al. [31] manufactured $Ti/TiO_2/Pt$ device where Ti/TiO_2 contact is ohmic and TiO_2/Pt is Schottky type. The negative voltage applied to platinum anode drifts oxygen vacancies towards TiO_2/Pt interface thus reducing the width of interface. Applying a

positive voltage would have an opposite effect. The same is true also in perosvkite materials as the charging effect at the interface also plays an important role [32].

There are some disadvantages for interfacial type switching compared to filament type switching. The biggest downside is the necessity of complex perovskite oxide that may not be compatible with CMOS manufacturing process. Also the noble metal platinum is needed for a good Schottky barrier. Interfacial type devices have thicker oxide material than filamentary type devices. This may have some problems in future when scaling becomes more relevant. However, there are also some advantages as asymmetric behaviour and non-linearity in $i - v$ plane allows operation of self-rectifying cell. Self-rectifying resistive switch acts itself as a diode thus removing the need for an external component. This is very beneficial in traditional memory arrays and also in 3D stacked memories [32]. Memory architectures using resistive switches are discussed more detail in Section 4.3.

3.4 Redox-Based Switching Mechanisms

3.4.1 Electrochemical Metallization Effect

In electrochemical metallization systems a conductive bridge (or channel) is formed between electrode materials where one electrode is electrochemically active metal and other electrode is electrochemically inert counter. Material between electrodes is a solid electrolyte. A very basic explanation of switching from HRS to LRS happens when active electrode metal, which is denoted as M, dissolutes according to, $\mathrm{M} \rightarrow \mathrm{M}^{a+} + ae^-$, where e^- is electron and a is positive charge in M^{a+} *cation*. Migrations of cations happen under high electric field in electrolyte material before an electrocrystallization occurs on the surface inert electrode, $\mathrm{M}^{a+} + ae^- \rightarrow \mathrm{M}$. The metal filament is broken when voltage that has reverse polarity is applied thus making the previously explained process to happen backwards. Some cations still remain in electrolyte so the device actually never returns to initial state like it was before the first forming step. This is the probable reason that initial forming voltage is higher than consecutive set voltages because no M^{a+} cations are present in fresh device. Some of the most typical electrode materials that are involved in ECM effect are Cu and Ag. Electrolyte materials such as SiO_2, Al_2O_3 or a-Si have large conductivity of the electrode cations. A very detailed analysis of fundamental processes behind electrochemical metallization memories was given by Valov et al. in [33].

ECM effect is also associated in literature with *programmable metallization cells* (PMC)

and especially *conductive bridge random access memories* (CBRAM). The term CBRAM has become a registered trademark but it is still used widely in academic literature. It can be argued that CBRAM and ReRAM are different technologies but in this thesis, all redox-based reactions are summarized under the term ReRAM.

3.4.2 Valence Change Memory Effect

The valence change systems also show bipolar behaviour although electrode materials do not insert cations into metal oxide. The polarity of switching is affected by many factors such as oxygen affinity of electrode or work function of metals. In a structure where mixed ionic-electrode layer is between active and ohmic electrode materials, a simple explanation of filament formation would be a mixture of oxygen vacancies and metal cations. The structure of such filament is quite even compared to electrochemically formed filament and is prone to have much higher resistivity [34]. Valence change systems are not only related to filament type switching but they are also present in interface type switching. Typically oxygen vacancies V_{O^-}, that are *anions*, have higher mobility than metal cations in many transition metal oxides. Higher electrical conductivity is achieved if the amount of oxygen vacancies are controlled during manufacturing process as it has an effect on valence state in transition metal cations. This results very fast switching times that several papers report to be less than 10 ns. For example, some TaO_x based devices that utilizes TaO_2 instead of fully oxidized Ta_2O_5 have much higher conductivity. In $Pt/TaO_x/Pt$ structure [35], bipolar switching happens when oxygen ions move near active electrode changing material stoichiometry from Ta_2O_5 to TaO_2 according to reaction $2TaO_2 + O^{2-} \leftrightharpoons Ta_2O_5 + 2e^-$.

3.4.3 Thermochemical Memory Effect

Previous two effects are observed in bipolar switching but unipolar switching is usually associated with thermochemical systems. In these systems, the resistive switching occurs when a filamentary thermal breakdown happens in oxide material. During the set process a weak filament is formed between electrodes because compliance current limits the formation. There is no current limitation present in reset process, and thus the filament gets broken by thermal processes. The soft breakdown that happens inside the oxide material by applying an electric field over device is induced by a thermal runaway mechanism. The localized thermal runaway under the weak electric field leads to a temporary on-resistance state that is known as a *threshold* switching. Increasing the electric field induces a redox

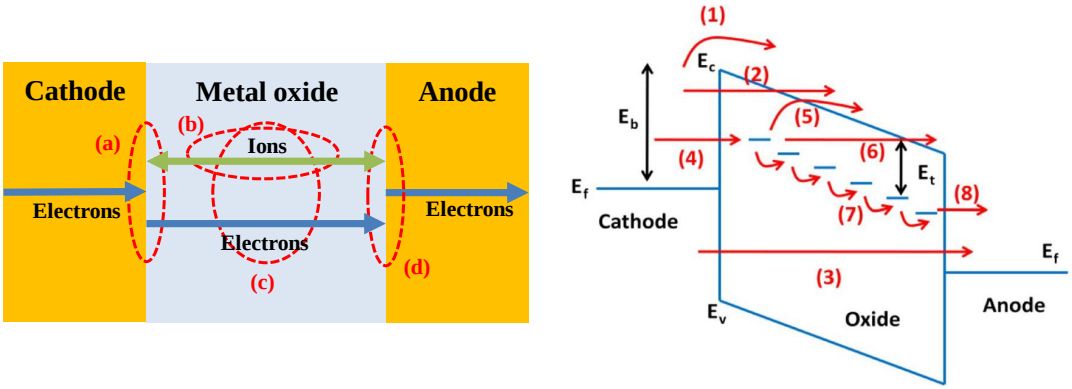

Figure 3.6: Relevant processes in redox based switching. **Left:** (a) anodic redox process (ECM), (b) electron or ion transport (VCM) (c) thermodiffusion and thermochemical redox process (TCM) (d) catodic redox process (ECM). Adapted from Waser's IEDM 2011 presentation. **Right:** The possible electron conduction paths of TiN/HfO_x/Pt metal-insulator-metal structure. (1) Schottky emission, (2) Fowler-Nordheim (F-N) tunneling, (3) direct tunneling, (4) tunneling from cathode to traps, (5) emission from trap to conduction band, also known as the Poole-Frenkel emission, (6) F-N like tunneling from trap to conduction band, (7) trap to trap tunneling, and (8) tunneling from traps to anode. In this figure the oxide material is n-type and E_c is the energy of conduction band, E_v is the energy if valence band, E_b is the bandgap energy, E_t is trap level energy and E_f is Fermi energy. Taken from [37].

based reaction because oxygen atoms drift away from high temperature regions as low temperature regions in transition metal oxides have lower valence states. Parts of metal oxide layer becomes metallic because they lack oxygen thus a conducting stable metallic type filament is formed. Filament rupture happens when thermochemical processes happen on the surface of filament. Because there is no current limitation involved in reset process, Joule heating effect forms a hot region somewhere in conducting filament which starts to dissolve. This process is self-accelerating because current density in filament starts to increase as filament itself is dissolving due to high temperatures. Process leads to complete dissolution and rupture of filament thus resetting device back to HRS [20].

The threshold voltage depends on thermal properties of electrodes and oxide materials and some examples of materials where thermochemical type behaviour is observed are NiO_x [36], and TiO_2 [19], between Pt electrodes.

Figure 3.6 shows conceivable redox-based processes in metal-insulator-metal stack. It also covers possible conduction paths in TiN/HfO_x/Pt device that was studied by Yu et al. [37]. The figure shows that while it is possible to cover all *relevant* processes behind switching the actual processes depend on the choice of materials. Some devices and their switching behaviours are presented in Table 3.1. It is important to understand that sometimes the redox-based process behind switching is not completely clear to authors as they

have not analysed their device using e.g. *in-situ* transmission electron microscopy.

Most common electrode materials are Pt, Cu or TiN while binary oxides such as AlO_x, CoO_x, CuO_x, HfO_x, NiO_x, SiO_x, TaO_x, TiO_x and WO_x have been studied thoroughly in academic literature. In memory technologies it is important to have high endurance, retention and low variability. Material selection criteria was first proposed by Yang et al. [38]. Still there is no material combination that has superior properties compared to rest. Rather than finding better or more exotic materials the most probable ways to improve devices switching properties are precise control of manufacturing process, purity of electrodes and correct oxygen doping of oxide material.

Table 3.1: Examples of different devices and their switching behaviour found in literature.

Device structure	Switching type	Redox-based switching	Ref.
$Cu/TaO_x/TiN$	Bipolar	ECM	[39]
$Cu/SiO_2/Pt$	Bipolar	ECM	[40]
$Pt/TaO_x/TiN$	Bipolar	VCM	[41]
$Ti/HfO_x/TiN$	Bipolar	VCM	[42]
$Ta/Ta_2O_5/TiN$	Bipolar	VCM	[43]
$Pt/TiO_2/Pt$	Bipolar, Unipolar	TCM	[2, 19]
$Cu/NiO_x/Pt$	Unipolar	TCM	[44]
$Pt/HfO_2/Pt$	Unipolar	TCM	[45]
$Pt/CoO_x/Pt$	Unipolar	TCM	[46]

In short conclusion of this chapter it is important to remember that the graphical explanations used in literature are only artistic approaches. There exist almost as many figures as there are authors. Figures are good illustrators but they do not replace mathematical equations. Instead, they should be treated more like an addition to mathematical models. It may be possible that very small resistive switches are affected by quantum mechanical phenomena. Complex quantum mechanics are very difficult to present using classical particle models.

4 Applications of Resistive Switching Devices

The industry and scientific community took interest after HP's publication. There were speculations about different ways to utilize these components from memory applications to learning neural networks. At the time of writing this thesis, a company called Adesto Technologies is the only operator in the market that has a ReRAM (CBRAM) based product. Companies like Sharp, Samsung and Intel besides HP are doing research on resistive switches aiming to commercialize them. In 2015, Intel announced their next generation 3D XPoint non-volatile memory architecture which was speculated to be a resistive switching based component that would hit the consumer market in 2016. The most research on resistive switches are towards a non-volatile ReRAM which is speculated to replace current industry standards like Flash-memory. Details on ReRAM memory and comparison to other memory types is presented Section 4.3.

Electronic devices and their applications can be roughly divided into two categories: *digital* and *analog*. The same division can be applied to resistive switching devices. In digital domain two or more discrete resistance levels are used between HRS and LRS. In analog domain continuous range of different resistance values between HRS and LRS are utilized. Figure 4.1 shows the division between the categories and some proposed application fields.

Academic publications of manufactured devices show promising results of desirable properties such as high-speed, low power consumption, non-volatility, n-state binary operation or even continuous analog operation although all of them may not be present at the same time. This could be an upgrade to current CMOS industry as further scaling of CMOS devices will become extensively difficult. Moore's Law predicted the scaling of transistor to become smaller every year and eventually the gate length of transistor will be in range of few nanometers which will be a big challenge as the length of conductive

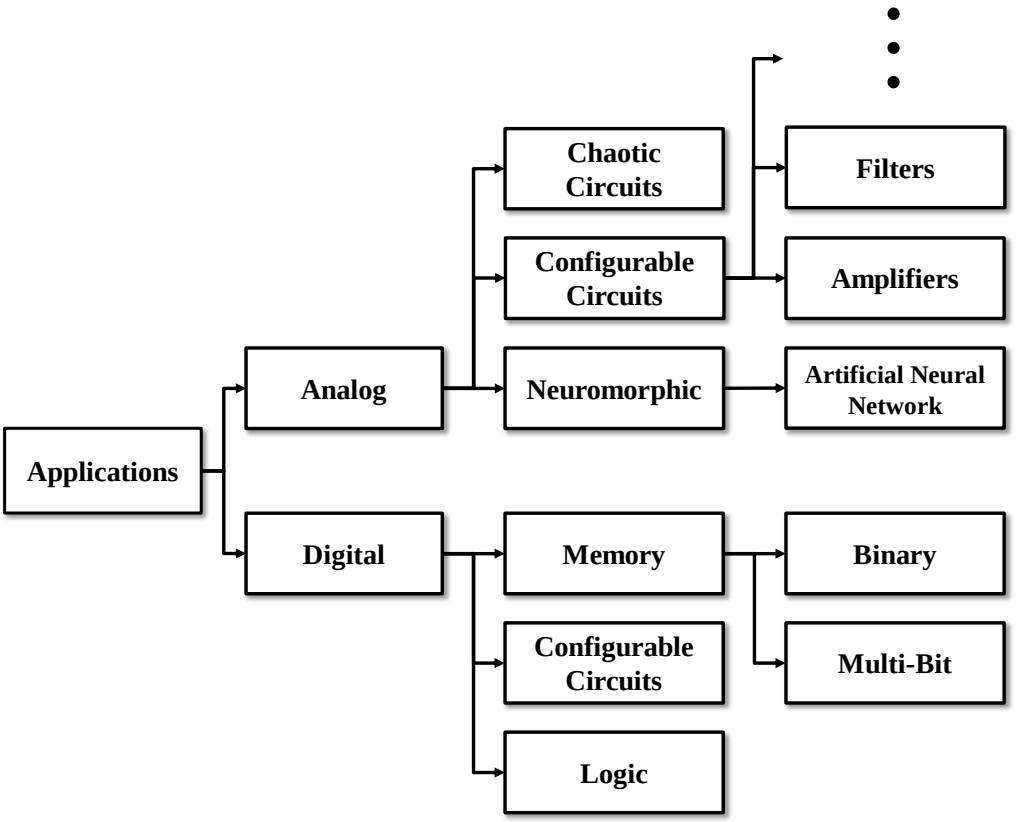

Figure 4.1: Examples of different proposed applications for resistive switching components. Application fields are divided into digital and analog domains based on the usage of device.

channel will be in same range as a single atom of silicon (van der Waals radius). Combining resistive switches with CMOS technology would at least give some extent to the lifetime of traditional CMOS technology. Building integrated circuits using only resistive switching components is not very convenient. Passive components can not provide any amplification or power supply to other components thus they have very limited options in pure computation. However, hybrid CMOS circuits are feasible to manufacture because some proposed materials used in resistive switches are compatible with standard CMOS industry. Manufacturing process should be simple because dimensions of resistive switches are straightforward to control. There are still many obstacles which prevent such circuits from being utilized in practice. Many times the manufactured devices suffer from great variability and repeatability of their behaviour is questionable. The situation gets even worse if statistical variation between devices is considered. The delays of commercial products seems to be affected by these constraints. For example, in 2014 HP Labs announced their future project "Machine" to be based on memristors but in 2015 they postponed the memristor technology in project roadmap.

4.1 Digital Applications

The most researched topics of proposed digital applications are digital memories, logic and reconfigurable circuits. The basic principle behind digital electronics is to decode voltage or current signal into logical values that are denoted as binary numbers $B = \{0, 1\}$. Single device can store either logical '1' or '0' depending on its resistance state, R_{ON} or R_{OFF}.

One of the scientific breakthroughs in 20th century was the invention of digital electronics by Claude Shannon in 1938 [47]. His work was prior to transistor and especially *metal-oxide-semiconductor field-effect-transistor* (MOSFET) that todays digital electronics heavily rely on. Shannon proved that it was possible to compute every Boolean function by implementing Boolean logic operators using relays and switches. In logic, Boolean function has a form $f : B^n \to B$, where $n \geq 0$ is the *arity* of the function. Basic operators in Boolean algebra are {AND, OR, NOT} that can be written as $\{\wedge, \vee, \neg\}$ notations. It is proven that subsets $\{\wedge, \neg\}$ and $\{\vee, \neg\}$ can synthesize any logical function in f thus the subsets are *computationally complete*. MOSFETs are used to build AND-, OR- and NOT-gates to implement logical operations in digital electronics, for instance. It is possible to construct other computationally complete logic basis using another operator named *material implication* that was proposed in the early 1900s by Whitehead and Russell in *Principia Mathematica* [48]. Material implication $p \to q$ ("if p then q" or "p implies q") is logically equivalent to $\neg p \vee q$. Operators $\to$ and FALSE (equivalent to logical '0') form a computationally complete logic basis. It turned out that implication logic was a natural way of constructing Boolean functions using memristors as Philip Kuekes proposed in 2008. Opposed to traditional digital logic gates, memristive logic is able to compute functions and at the same time store information thus acting as digital latches. The resistive switches are used in *stateful logic* where the resulting state depends on the value of the previous state. Different approach to compute logical functions must be taken because fundamentally memristors are two-terminal passive components and they cannot take two inputs simultaneously which is needed in AND- and OR-gates. Experimental result were obtained in [49], where two or three $\mathrm{Pt/TiO_2/Pt}$ based memristors were connected in parallel to perform simple Boolean functions. It is proven that two components are enough to compute all Boolean functions [50], though intermediate results must be stored in computations. Interesting approach is to compute logical operations and store data in the same memory space instead of traditional Von-Neumann or Harvard architecture.

Another field of using resistive switches is digital reconfigurable logic. Especially *field-programmable-gate-arrays* (FPGA) are potentially easy to manufacture with hybrid CMOS technology. FPGA is a digital circuit which can be reprogrammed to perform different computational operations. Big part of FPGA chip is reserved for *look-up table* memory that stores configuration bits and only small proportion is used for actual computations. Resistive switches could provide a huge area benefit by storing the routing bits between computation logic blocks and leaving actual computation inside computation blocks for CMOS [51].

There are many types of different proposed architectures to combine resistive switches and CMOS. In [52], Strukov proposed a CMOS and molecular scale nano-switches (CMOL) architecture that combines traditional FPGA-type architecture with parallel nanowires that are used for switching. This architecture is known as CMOL-FPGA.

4.2 Analog Applications

The approach to utilize resistive switches in digital applications is not very innovative as there already exists multiple different technologies that can be used in memories or computations efficiently. However, if the continuous resistance range between R_{ON} and R_{OFF} could be utilized, the application field of resistive switches would become much wider. Resistive switches could provide simultaneously big performance boost, lower power consumption and area efficiency that are difficult to achieve with currently available technologies.

Pershin and Di Ventra demonstrated the possibilities to use resistive switches in analog applications such as comparators, gain amplifiers and oscillators [53]. For example, a comparator's operation is based on a threshold voltage which is compared to input voltage and the output is based on a decision on which one of the two voltages is higher. Typically, a comparator has a fixed threshold voltage that is set with e.g. resistors. A configurable resistor such as potentiometer is difficult to built efficiently on silicon. Instead, a resistive switch component could produce multiple threshold voltages for a comparator. Another circuit example is a programmable gain amplifier. A gain amplifier is an operational amplifier whose gain is decided on a feedback loop. In a non-inverting case the input signal is applied to the positive input of operational amplifier and negative input level is determined by resistor based voltage divider circuit. A simple voltage divider gives a

gain that is proportional of the ratio of the two fixed value resistors. However, if one of the resistance values could be modified the gain of the amplifier would also increase or decrease. A resistive switch would act as a programmable component thus modifying the gain of the amplifier circuit. The resistive switch would essentially be a type of digital potentiometer rather than fully analog component in previously proposed application. However, it would still provide a huge benefit in area efficiency compared to transistor-based digital configurable circuits in this particular use case.

Theory of utilizing resistive switches is not limited to basic analog building blocks. In 1983, during Chua's visit to Waseda university, he suggested a circuit [54], that consists of two resistors, two capacitors, a negative resistor, an inductor and a non-linear resistor. This theoretical circuit is known as *Chua's circuit* or *Chua's oscillator*. Chua's later work improved the circuit model equations to work with memristor [55]. The dynamical and highly non-linear operation can be used to build chaotic circuits that have applications in cryptography, secure communication or even medical purposes [56, 57, 58].

Finally, perhaps the most promising applications in analog domain are *neuromorphic circuits*. Basically neuromorphic circuits try to mimic the biological neural activity of human or any other animal. Biological neural networks are cells that carry information through electrical or chemical signals. The cells in such systems are called *neurons* and signaling between neurons happens via *synapses*. Implementation of a neuromorphic system in circuit level can be done by using previously mentioned CMOL-type structures, where CMOS components act like neurons while synapses are resistive switches built on top of CMOS-layer in a net like structure. Two implementations of neuromorphic architectures are *artificial neural networks* (ANN) and *cellular neural networks* (CNN). Artificial neural networks are discussed later in Section 4.4 so a brief explanation of CNN is given here.

Cellular neural networks differs from biological neural networks in a way that communication is allowed only between the nearest neighbouring cells. CNN was first proposed by Chua and Yang [59], to be a circuit that parallel processes analog signals in real time. Cells in CNN are circuit clones of each other but the functionality of cells can be different during computations. The resistive switch implementation of CNN consist of a non-linear CMOS computational cell and a resistive switch network. This implementation provides an area efficient solution to conventional parallel computation structures because the inter-cellular communication is lifted above the CMOS circuit. A two dimensional network is a construction of rectangular $M \times N$ array, where M is number of rows and N is number of

columns. This type of structure has become a standard CNN architecture [60]. Originally cellular neural networks were proposed to solve image processing and pattern recognition problems [59].

4.3 Resistive Random Access Memory

The most research to utilize resistive switching components goes towards resistive random access memories. ReRAM is only one of many existing or emerging technologies such as static random access memories (SRAM), dynamic random access memories (DRAM), Flash, spin-torque transfer RAM (STT-RAM), magnetic RAM (MRAM), phase change memory (PCM), hard disk drive (HDD), for instance. Memory types can be categorized into two different categories: *volatile memories* that require power to store the information and *non-volatile memories* (NVM) that remember their stored information even if power is temporarily cut-off. ReRAM is a non-volatile memory and it is natural to compare its performance to other non-volatile technologies like Flash. Based on performance requirements of current NVM, some performance parameters that ReRAM must satisfy to compete with existing technologies are listed below. These parameters also serve as performance metrics for the manufactured ReRAM device of this study:

Write and Read Operations

Flash memories have very high programming voltages, typically over 5 V and erase voltages in range of 10 V. Typical Flash programming speed is around 10 µs. Reported resistive switching components have write voltages less than 5 V, sometimes in the range of few hundred millivolts. High programming voltages are not a problem because power consumption can be limited with compliance current. However, read voltages in ReRAM are much lower than Flash. Less than 100 mV is enough to give proper read current for detection sense amplifiers. Switching time in ReRAM is very fast. Preferably it should be less than 100 ns to give a big benefit margin over Flash. Some TaO_x-based components have programming time < 10 ns which makes it comparable with DRAM [61]. However, the reset process takes a time that is many orders of magnitude longer than the set process.

Endurance

The maximum number of write cycles in Flash depends on its type, but NAND-type Flash has typical endurance of 10^3 to 10^5 cycles. Some specific algorithms can be used in Flash memories to increase write endurance. ReRAM devices also have very wide range of endurance cycles depending on the used materials. ReRAM should endure at least 10^6 cycles but preferably many orders of magnitude more if computational operations in ReRAM are also considered.

Retention

Commercial NVM products usually guarantee data retention time to be more than 10 years, which should be the minimum requirement for ReRAM. Retention must be preserved with thermal stress conditions up to 85 °C and electrical stress under continuous read cycles. Typically, ReRAM devices suffer from gradual loss of data retention where $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ decreases over time. At the time of writing this thesis it is still too early to say if data retention time up to 10 years is possible because most measurements have been started after 2008, although accelerated tests are possible. Studies have reported devices that can at least store information longer than year thus ReRAM could be considered as a proper NVM technology.

Multi-Bit Operation and Scaling

Flash memory cell which can store multiple bits is called *multi-level cell* (MLC). MLC is beneficial because it leads to higher data density but the trade-off is worse endurance. ReRAM has also potential to act as a multi-bit storage memory as is mentioned in the beginning of this chapter. However, such behaviour requires good control of resistance values between R_{OFF} and R_{ON}. A unique property of ReRAM is its on-off ratio which should be large enough to fulfil necessary requirements for NVM. In practice, ReRAM might need $R_{\mathrm{OFF}}/R_{\mathrm{ON}} > 10$ for reliable operation. Smaller value is also acceptable if considered application is very specific. Otherwise reliable operation becomes an issue when e.g. data retention is considered. Device variability is also a problem in ReRAM devices. Bigger $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ could provide a safe operation margin even if statistical variability of resistance values is large.

The scaling of memory architectures has historically decreased over time. A geometry

scaling of memory technology is described as a minimum feature size F on a given process. If an area of memory matrix is A_{mem}, the area factor n tells how many squares of feature size F^2 is needed to realize the cell

$$A_{\text{mem}} = nF^2. \tag{4.1}$$

In the case of Flash $n \leq 4$, while for some other memory technology such as DRAM the area factor is $n \approx 8$. Proposed ReRAM architectures have memory area $4F^2$ if a *crossbar array* is used. The term crossbar array is explained later but a short comparison of different storage technologies is given before that. Yang et al. gathered [62], data from International Technology Roadmap for Semiconductors (ITRS) to show how ReRAM fits into the current map of different memory technologies. Table 4.1 shows an adaptation from their paper. This table shows that ReRAM has many advantages compared to other NVM, especially Flash. Even though the energy that is needed to store one bit in Flash is negligible, the access to Flash cell needs >100 pJ thus making the operation energy many times higher. If operating speed is considered, ReRAM competes with DRAM. Combined with lower energy consumption and better endurance it could be a viable solution at least to partially replace DRAM in the future. PCRAM and STT-RAM are shown for comparison as they are also prototype NVM.

ReRAM memory matrix is proposed to be a crossbar type array where bit lines and word lines are perpendicular to each other. It is the most efficient way to manufacture a ReRAM array. The oxide material is manufactured into the cross section between metal

Table 4.1: Comparison of memory and storage technologies. SRAM and DRAM are the only volatile memories in this table.

	Density ($1/F^2$)	Energy/bit (pJ)	Read time (ns)	Write time (ns)	Retention (s)	Endurance (cycles)
ReRAM	4	0.1–3	< 10	< 100	$> 10^8$	10^{12}
Flash	≤ 4	0.00002	10^5	10^5	$> 10^8$	10^3–10^5
PCRAM	4–16	2–25	10–50	50–500	$> 10^8$	10^9
STT-RAM	20–60	0.1–2.5	10–35	10–90	$> 10^8$	10^{15}
SRAM	140	0.0005	0.1–0.3	0.1–0.3	-	$> 10^{16}$
DRAM	6–12	0.005	10	10	< 1	$> 10^{16}$
HDD	2/3	10^9	$\approx 10^6$	$\approx 10^6$	$> 10^8$	10^4

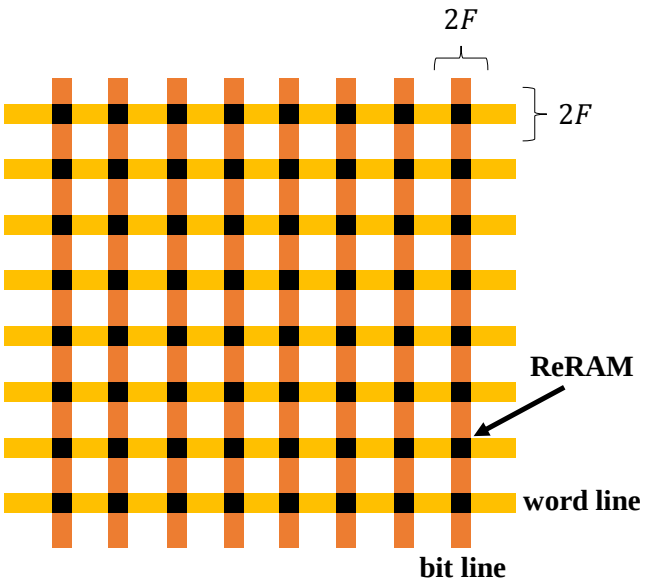

Figure 4.2: Crossbar structure of ReRAM array.

lines because a ReRAM device is fundamentally a metal-insulator-metal structure. A crossbar array structure was first proposed by Baek in 2005 [63], for an efficient way to construct memory arrays from resistive switching components. Figure 4.2 shows how the minimum area $4F^2$ of ReRAM array is achieved using crossbar structure.

The simple approach to construct a ReRAM array only from resistive switching components has its limitations. Components have undesirably high non-linearity and sometimes limited $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ or very low R_{ON}. Analytical approaches to simulate crossbars showed that previously mentioned constraints limit the array size considerably [64, 65]. Perhaps the biggest problem is the *leakage current sneak path* through adjacent components near a selected ReRAM device if passive crossbar structure without any external components is used. Leakage current increases power consumption and causes read errors. Resistive switching component itself has to be rectifying, otherwise the current would flow through all devices in the array. Diode like operation can be achieved if interfacial-type component that uses perovskite oxides is utilized. It is not a perfect solution because of the CMOS manufacturing constraints that are already mentioned in Section 3.3.

One of the most promising solution to solve current sneak path problem is to manufacture a MOSFET in series with ReRAM device. A MOSFET does not limit the choice of electrode and oxide materials of ReRAM. It also acts as an isolator, a selector and most importantly a current limiter. A structure where MOSFET is combined with ReRAM

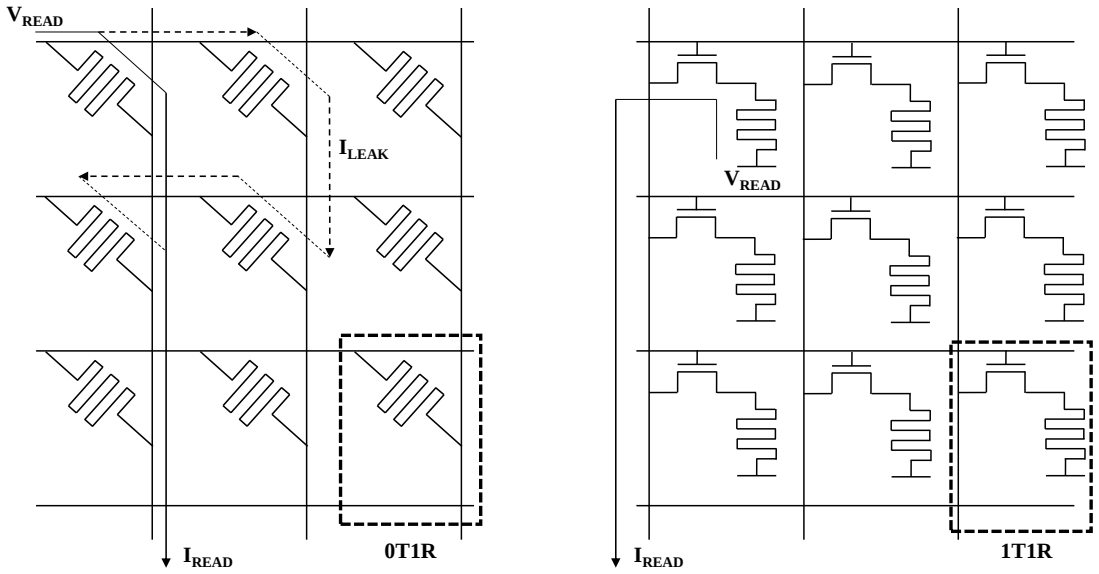

Figure 4.3: 0T1R and 1T1R structures of ReRAM array. The leakage current goes through half-selected cells without a MOSFET selector.

is called 1T1R whereas passive structure is called 0T1R. These structures are shown in Figure 4.3 that also shows how leakage current finds its path through adjacent devices in a passive array. 1T1R structure has also limitations because transistors must be built under ReRAM array which limits the stackability and 3D ReRAM array options. MOSFET also limits the ReRAM array size.

Another way to solve the sneak path problem is to use an external diode (1D1R). Diode is a good selector because it has a small size and at the same time it is a two terminal passive device that can easily be stacked. Like any other external component solution, a diode has its trade-offs. A simple diode can be used only with unipolar devices and it has a relatively large voltage drop that limits the low programming voltages of ReRAM.

4.4 A Brief Introduction to Artificial Neural Networks

Artificial neural networks (ANN) try to process information and perform computations in a similar topology than human brains. Problems like facial recognition, speech recognition or language processing are considered hard to implement in traditional computer algorithms while those tasks are trivial to humans. However, imitating the behaviour or the structure of the brain is not a simple task[1]. Research about ANN has been going on for decades and there have been successful attempts in software level for e.g. recognizing objects [66], and

[1]Human brain has approximately 10^{11} neurons each with around 1 000 synaptic connections.

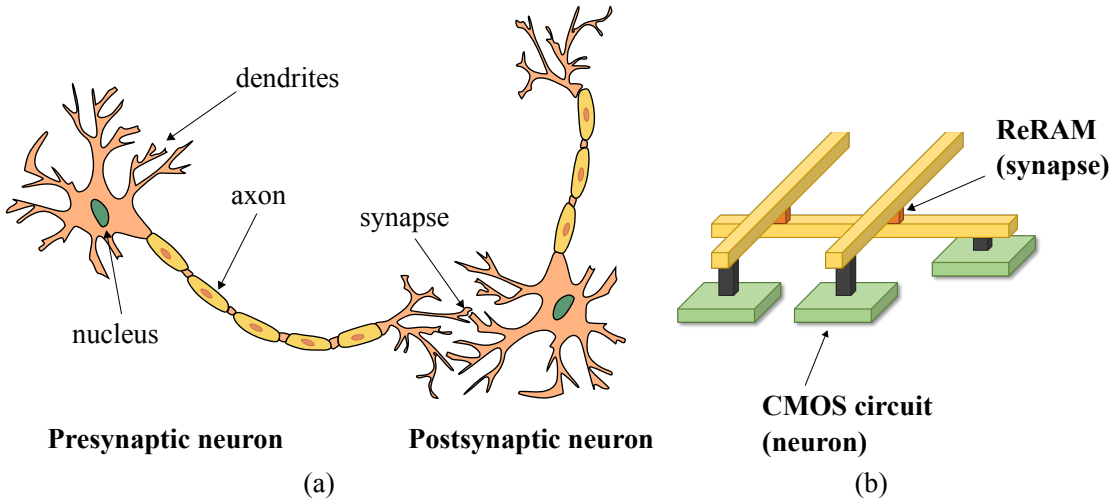

Figure 4.4: (a) A biological neural network. (b) Neural network implemented on harware using crossbar ReRAM array.

speech [67]. Also, a strong artificial intelligence with deep learning has been successfully implemented to play the game of Go [68]. Achieving ANN on software is not very efficient and cost effective because typically big computer clusters or even supercomputers are needed to compute multiple complex parallel computations. Truly parallel processing capabilities are efficiently achieved using hardware solutions. Misra et al. gave a great review [69], how ANN have been realized on hardware level in recent years.

It turns out that crossbar ReRAM array is a very natural way to construct ANN. Figure 4.4 shows the parts in biological neural network and the possible hardware implementation using ReRAM. The information inside neural network flows in one way from pre-synaptic neuron to post-synaptic neuron.

Mathematically a neural network can be presented in multiple ways. The network can be recurrent with feedback loops and it can be shown that recurrent networks are universal computers where feedforward networks are not. A feedforward network is a system whose inputs go through multiple layers of computation called *hidden layers* before outputs are calculated. One way to present such a system mathematically is

$$b_j = \sigma\left(\sum_i a_i w_{i,j}\right), \tag{4.2}$$

where b_j is the output, a_i is the input, σ is some non-linear function and $w_{i,j}$ is the synaptic

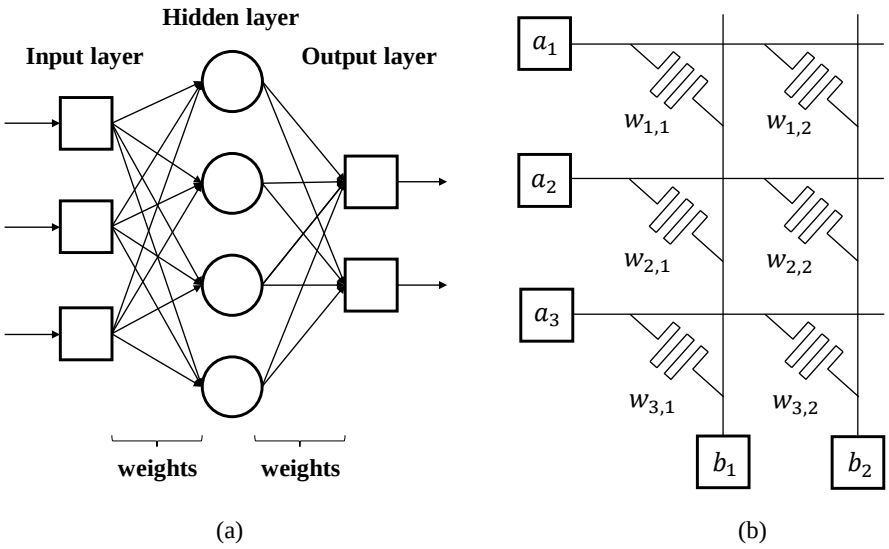

Figure 4.5: (a) A diagram of feed-forward network of ANN with three different layers whose input nodes are multiplied with weights and summed together. (b) ANN made of 2×3 crossbar ReRAM array.

weight. A typical graphical representation of feedforward ANN is shown in Figure 4.5 (a) where inputs are fed through a hidden layer before outputs are given. The number of nodes should be optimized to work with a given problem. An example of 2×3 crossbar ReRAM array is also shown in Figure 4.5 (b). It can be seen that a matrix multiplication that (4.2) represents is a natural way to present ReRAM matrix in analog manner. The analog input voltages $\{a_1, a_2, a_3\}$ are multiplied with pre-programmed resistance values $w_{i,j}$, where $j = 1, 2$ and $i = 1, 2, 3$. The output currents $\{b_1, b_2\}$ are calculated by summing the products of inputs and weights. It is also possible to perform further computation to outputs by running them trough e.g. analog/digital filter which would represent the non-linear σ in (4.2). The limiting performance factors of such ANN are the resistance values of ReRAM devices. Without good analog or multi-bit operation the ANN is limited to perform tasks in binary domain.

Attempts to manufacture analog ReRAM devices to be artificial synapses have been made. Fully analog operation is difficult to achieve because there exists variability between switching cycles. Pseudo-analog operation where multiple discrete resistance levels between R_{OFF} and R_{ON} are used, is a more realistic approach. There is no exact number of different resistance levels that ReRAM in ANN should have. The objective is to have as many as possible repeatable resistance values. Some research groups have demonstrated a pseudo-analog operation [70, 71], or good control of multiple resistance levels [72], for neuromorphic computing purposes.

In this thesis the proposed device for artificial synapse is Ta_2O_5-based ReRAM. In Chapter 6 and Chapter 7, the suitability as well as general characteristics of the device are analysed. Previous studies of neuromorphic ReRAM devices have used HfO_x [72], TiO_x [70], or TaO_x [73], oxide materials, so Ta_2O_5 should be a good candidate.

5 Fabrication of $Ta/Ta_2O_5/TiN$ ReRAM Devices

The initial plan was to fabricate four different $Ta/Ta_2O_5/TiN$ ReRAM cells with varying oxide thicknesses, $t_{ox} = \{5\,nm,\ 7\,nm,\ 10\,nm,\ 20\,nm\}$. The purpose and choice of such oxide thickness and tantalum top electrode material was to have comparison data with previously manufactured $Cu/Ta_2O_5/TiN$ ReRAM cell that had same oxide thicknesses but different top electrode material manufactured. As explained later in this Chapter, due to fabrication issues the oxide thickness of 5 nm did not yield functioning devices. Initially the samples were cut from a silicon wafer to be 20 mm × 20 mm in size. The fresh substrate samples were provided by Kyushu Institute of Technology. A 200 nm thick SiO_2 layer was initially deposited on top of substrate.

Chips were manufactured at Hokkaido University's Laboratory of Nanomaterial Science.

5.1 Organic and Inorganic Cleaning

The purpose of organic and inorganic cleaning is to remove any impurities and organic contaminants from the substrate before further processing steps. Inorganic cleaning is sometimes called as *ionic cleaning* because it's purpose is to remove ionic contaminants from the sample.

In organic cleaning, all four 20 mm × 20 mm samples were cleaned with 99.5% ethanol and acetone. The first step was to soak samples in acetone. Samples were put inside ultra sonic automatic washer, where they were cleaned for five minutes. This step was repeated before samples were washed using ethanol. Finally the samples were dried out using nitrogen gas.

In inorganic cleaning step sulphuric acid, H_2SO_4, and hydrogen peroxide, H_2O_2, were

used with ratio of 3:1. Samples were cleaned for five minutes in this solution. Finally the samples were washed in deionized water using ultra sonic automatic washer.

5.2 Deposition of Bottom Electrode and Insulation Layer with Sputtering

The next manufacturing step was to deposit bottom electrode material and silicon oxide layer. Bottom electrode materials were deposited on substrate by using a radio frequency sputtering method (RF sputtering) at room temperature. Depositing thin-film structures using this sputtering method involves a target material that provides source materials for substrate. The first sputtering phase included three targets: Ti, Pt and SiO_2. The RF sputtering uses magnetrons to create strong magnetic and electric fields in vacuum chamber. The sputtering requires a high-density plasma and it is necessary to perform at high vacuum conditions in order to facilitate the deposition. Magnetic field forces ionized gas particles to collide with target material. Typically inert sputter gas such as argon is used. A collision of charged Ar^+ high kinetic energy particles forces target material atoms to emit into surrounding space. These particles are not electrically charged so they do not react with magnetic or electric fields thus flying in straight lines. Some portion of these particles hit the substrate. The sputtering machine used this study was the high-frequency magnetron sputtering device E-200S manufactured by Canon Anelva.

Figure 5.1 shows the sample after material deposition. The bottom electrode material in this study was TiN but previous experiments using this material showed that parasitic resistance of TiN material was too high for reliable switching behaviour. To reduce the effect of parasitic resistance, an extra Pt layer was deposited between Ti thin-films. Ti films prevent material peel off from SiO_2. The SiO_2 layer is used in photolithography step to allow bottom electrode patterns to be manufactured.

Parameters for sputtering were obtained from previous experiments of ReRAM manufacturing attempts. These parameters are shown in Table 5.1. The inert gas argon was used as main sputtering gas. Deposition of TiN layer needed nitrogen gas where 80:20 partial ratio was used between Ar and N_2.

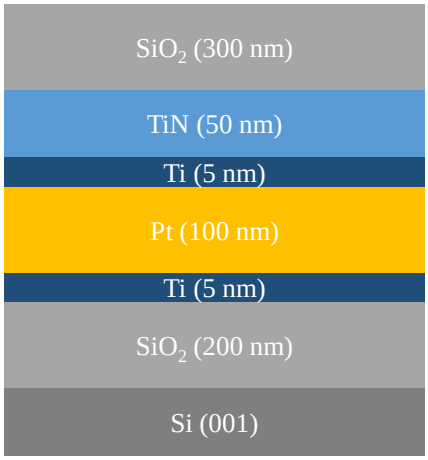

Figure 5.1: Sample after deposition of bottom electrode materials.

Table 5.1: Sputtering parameters used in bottom electrode deposition. [1]Sputtering gas ratio ($Ar:N_2 = 93:7$).

Target	Gas pressure (Pa)	RF power (W)	Target distance (mm)	Sputtering rate (nm/min)
SiO_2	1.00	150	95.0	5.0
TiN[1]	1.00	100	80.0	2.1
Pt	1.00	100	95.0	10.0
Ti	1.00	100	80.0	4.1

5.3 First Photoresist Application and Photolithography

Photolithography is used to transfer patterns onto sample for further processing and material deposition. A light sensitive *photoresist* application is applied on top of the substrate. An even distribution of photoresist is necessary to have correct pattern formation which is achieved by *spin coating*. The samples are rotated at high speed during application. Fluid-mechanisms ensures that top of the photoresist layer moves faster than the layer near the substrate surface thus producing an uniform layer of photoresist. After spinning the substrate is prebaked to remove excessive solvents of photoresist.

Next step is to expose substrate under intensive ultra-violet light. UV-light causes chemical changes in photoresist which makes it possible to remove exposed resist using chemical that is called a developer. Because the purpose of photolithography is to transfer patterns onto substrate it is necessary to prevent UV-light from exposing the whole photoresist. A *metal mask* that has a desired pattern is placed between substrate and

UV-light during the exposure.

Finally the exposed parts of photoresist are removed in developing step. Development chemical reacts with the resist thus dissolving it. The remnants of unexposed photoresist is hardened by baking the substrate. Hardened photoresist acts as protective layer in next processing steps such as *wet chemical etching* or *reactive ion etching* (RIE).

In this study the first photolithographic step was used to transfer bottom electrode pattern onto the samples. The metal mask had also circle patterns for ReRAM device holes with following diameters,

$$\phi = \{2\,\mu\text{m}, 4\,\mu\text{m}, 8\,\mu\text{m}, 16\,\mu\text{m}, 32\,\mu\text{m}, 64\,\mu\text{m}, 100\,\mu\text{m}, 200\,\mu\text{m}\}$$

Initially *hexamethyldisilazane* (HMDS) was applied on samples to improve the adhesion between the substrate surface and actual photoresist, so that the photoresist would not peel off in development process. The spin coating device was Opticoat Spincoater MS-A150. The rotating time and speed settings are shown in Table 5.2.

Table 5.2: Spin coater settings.

Time (s)	Rotation speed (RPM)
0 - 3	0 - 300
3 - 10	300 - 3000
10 - 33	3000
33 - 40	3000 - 0

Samples were pre-baked at 110 °C for 90 s after first spin coating. The actual photoresist OFPR-800-34C was applied after pre-baking. Samples were placed again on the spin coater which used same setting as in previous round. After spin coating, the samples were baked again at 110 °C for 90 s.

The next phase was to expose photoresist under UV-light with protecting metal mask. Photolithography device was Suss Microtec MA6/BA6 Contact Aligner. The mask alignment was done before the exposure. Exposure time was calculated to be about 7.3 s based on the UV-light intensity measurement. Exposure time is equivalent to dose of 120 mJ.

After the exposure, development is performed for 90 s by using NMD-3 developer. To guarantee full development samples are examined under microscope afterwards. Finally the samples were post-baked for five minutes at 120 °C.

5.4 Reactive Ion Etching and O_2 Plasma Ashing without Tunnel

Reactive ion etching (RIE) is one type of dry etching used in microfabrication process to remove materials from the wafer by utilizing chemically reactive plasma. The etching device used in this study was EIS-700 manufactured by Elionix. The etching device consist of a chamber where a strong radio-frequency (RF) field is generated between a coiled antenna and a dielectric bottom electrode. The RF field is oscillating at 13.56 MHz which ionizes the gas thus creating plasma inside vacuum chamber. Inductively coupled RIE can be used to lower the ion energy, while increasing the plasma density. EIS-700 utilises *inductively coupled plasma* (ICP). It is possible to increase the plasma density in the vicinity of the electrode surface by combining magnetic field and the RF field. However, it is difficult to produce a uniform magnetic field in a wide plane which may lead to non-uniform plasma distribution in the substrate.

In previous step the areas of SiO_2 for bottom electrode contacts and ReRAM device holes were exposed for etching. RIE etching removes SiO_2 completely where photoresist does not exist thus revealing TiN electrode material. High electron density approximately $10^{12}/\text{cm}^3$ is obtained in range of 0.05 Pa to 1.00 Pa. The EIS-700 chamber was vacuumed to 0.17 Pa. Reactive gases used were O_2 and CF_4 in ratio of 2.9:8. Powers applied to antenna and electrode were 100 W and 20 W consecutively. Total etching time was seven minutes. The cross section of device is shown in Figure 5.2.

After etching it was clear that one sample did not etch perfectly. It may be possible that photoresistor development after exposure was not perfect and remains of photoresist existed on top of the sample preventing complete removal of SiO_2. Another possible reason for this was deficient exposure under UV-light. The thickness of photoresist application was probably different compared to the other samples. This could have been avoided by observing the sample under microscope more carefully after development. It may be necessary to dissolve the whole photoresist and start the lithography process from the beginning before RIE. The sample is shown in Figure 5.3.

The initial plan was to manufacture four chips varying Ta_2O_5 thickness, where $t_{ox} = \{5\,\text{nm}, 7\,\text{nm}, 10\,\text{nm}, 20\,\text{nm}\}$. It was decided that 5 nm oxide thickness was left out from this study because of this failed chip.

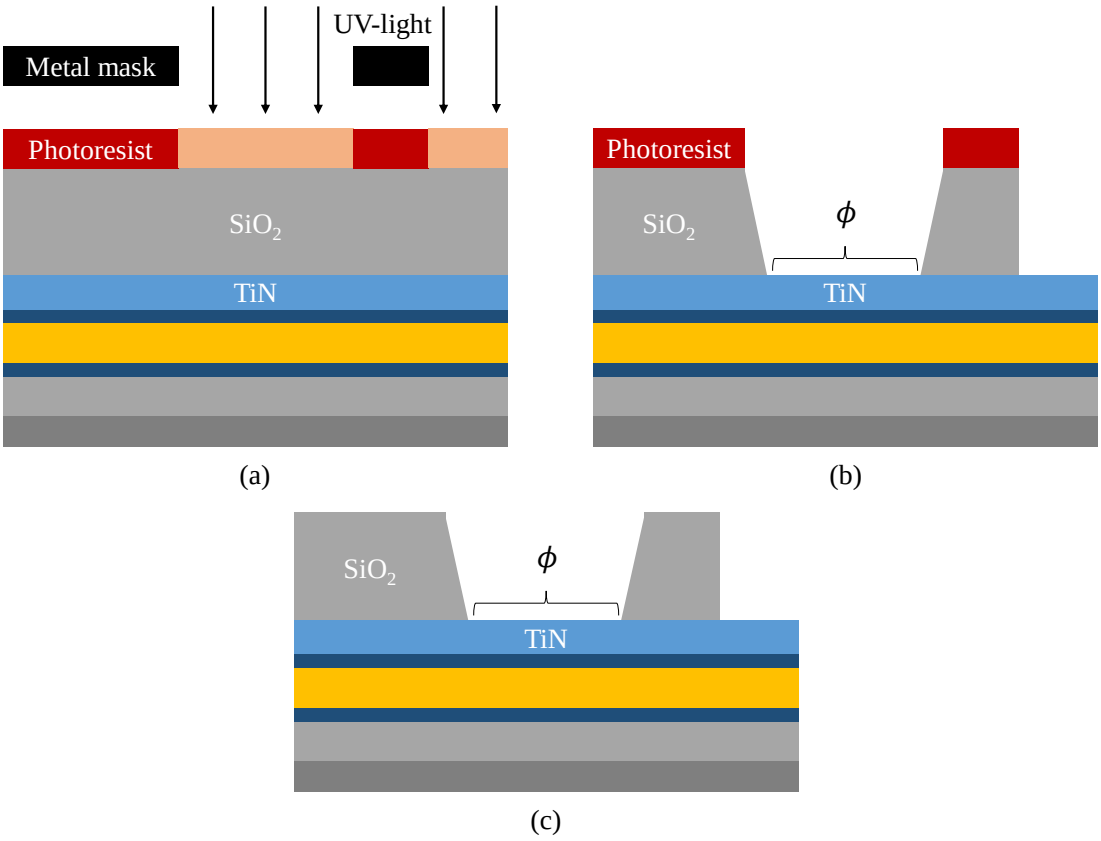

Figure 5.2: (a) Exposuring sample under UV-light after photoresist application. (b) Sample after RIE and (c) after O_2 plasma ashing.

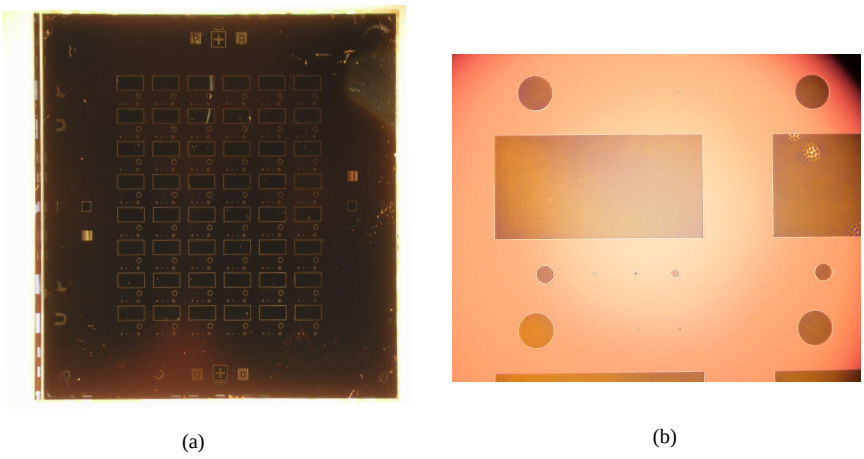

Figure 5.3: (a) Incomplete etching. TiN bottom electrode is not revealed because SiO_2 was not completely etched. (b) Magnified image of the same chip.

Before second photolithography step, the remaining photoresistor must be removed. Samples were cleaned in acetone and ethanol. This organic cleaning step was exactly the same that is explained in Section 5.1. However, organic dissolving is not enough to remove all photoresist and remainders are removed using a method called *plasma ashing*. In this

study O_2 reactive gas was used. Oxygen plasma was generated in vacuum chamber where O_2 gas was exposed to high power RF-field. O_2 plasma reacts with organic photoresistor material on the sample's surface to form CO, CO_2, and H_2O. This ash is removed by pumping it outside the vacuum chamber.

Yamato Plasma Reactor 500 was used to perform plasma ashing. The samples were placed inside the chamber *without* using a tunnel. When tunnel is not used the ashing process is more powerful thus removing all remaining photoresist. Vacuuming was performed to create pressure $< 10\,\text{Pa}$ inside the chamber. O_2 gas flow was set to 150 ml/min and RF power was set to 300 W. The ashing time was 10 min.

5.5 Second Photoresist Application, Photolithography and O_2 Plasma Ashing with Tunnel

The purpose of second photolithography was to transfer top electrode pattern to the samples. This step was very similar to the first photolithography step. Careful alignment is important to match ReRAM holes and top electrode patterns. Two photoresist applications were used. If only one photoresist application is used, the lift off phase may not be complete. Two types of different photoresist form a canopy like structure.

The first application was *polymethylglutarimide* (PMGI) and it was applied using the spin coater with the same settings that were explained in Section 5.3. After coating, samples were placed for 5 min on hot plate where temperature was 190 °C. Samples were let to cool down sufficiently after baking before the second photoresist was applied. The second resistor application, OFPR-800, which is the actual photoresist was applied by spincoating. Samples were placed on a hot plate and baked for 90 s at 110 °C after coating.

The next phase was to align samples and calculate exposure time that was equivalent to 120 mJ. Afterwards, development was performed for 200 s using the NMD-3 and rinsed after three times for 30 s with deionized water. Sometimes the single-layer resist remains or it is not fully developed so examination under microscope is necessary.

Before sputtering, plasma ashing was performed on samples with the same settings that are explained in Section 5.4. This time the samples were placed *inside* a tunnel during ashing to prevent complete removal of PMGI which is necessary for lift-off.

5.6 Deposition of Top Electrode with Sputtering and Lift-Off

After plasma ashing all samples were ready for top electrode deposition. A stoichiometric Ta_2O_5 target was used to deposit the ReRAM oxide layer. In this study, partial pressure of sputtering gases Ar and O_2 was decided to be 50:50 but other ratios could be investigated in further studies. Platinum is used for protective layer and to prevent peel-off.

Table 5.3: Sputtering parameters. Different oxide thicknesses, $t_{ox} = \{7\,\text{nm}, 10\,\text{nm}, 20\,\text{nm}\}$ were manufactured varying sputtering time calculated from sputtering rate value. [1]Sputtering gas ratio ($Ar:O_2 = 50:50$).

Target	Gas pressure (Pa)	RF power (W)	Target distance (mm)	Sputtering rate (nm/min)
Pt	1.00	100	95.0	10.00
Ta	1.00	100	80.0	8.10
Ta_2O_5[1]	1.00	100	95.0	1.28

The lift-off phase was the final step of manufacturing process. Remains of photoresist and previously sputtered material exists on top of the TiN bottom electrode. Process was similar to organic cleaning. Samples were initially cleaned with dimethylformamide (N,N-DMF) using ultra sonic automated washer to remove PMGI. Cleaning time was about five minutes but in some cases it was necessary to use longer times because PMGI was not completely dissolved. After all the PMGI was dissolved, samples were placed another N,N-DMF solution for another 5 min. Next step was to clean samples for 5 min using acetone. This step was repeated twice. Finally, the samples were cleaned in 99.5 % ethanol again for 5 min and dried using nitrogen gas. The cross section of device is shown in Figure 5.4 and the final manufactured chip is shown in Figure 5.5.

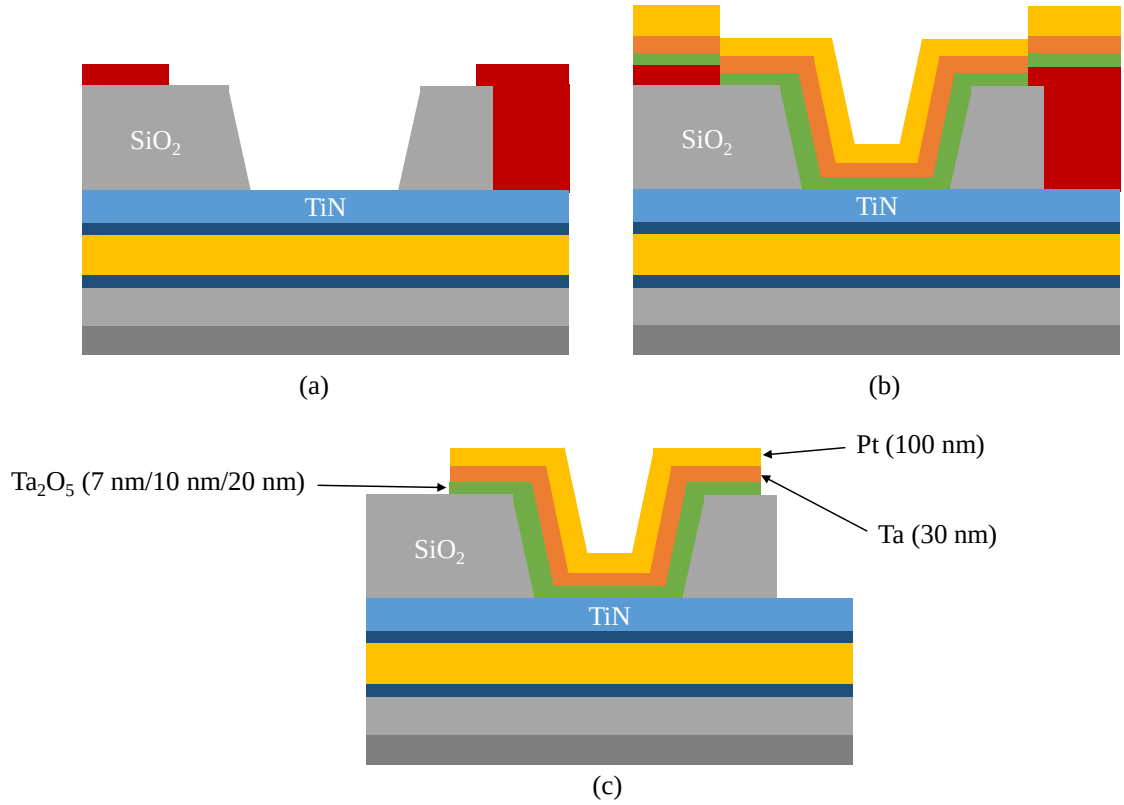

Figure 5.4: (a) Sample after second photolithography step. (b) Deposition of oxide and top electrode materials. (c) Complete device after lift-off.

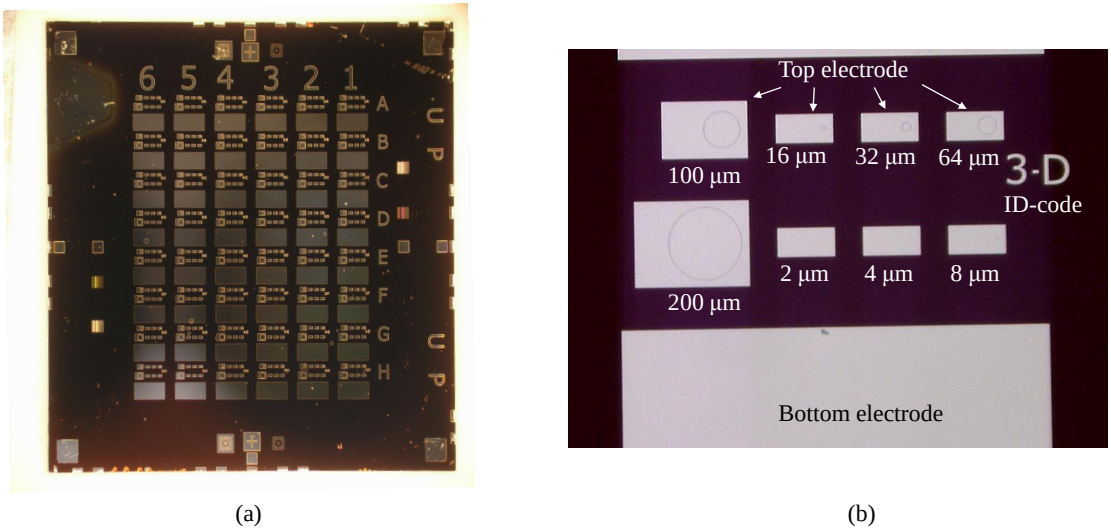

Figure 5.5: (a) One of the manufactured chips which is divided into 48 identical sections, $\{1, 2, 3, 4, 5, 6\} \times \{A, B, C, D, E, F, G, H\}$. (b) Magnified image of 3-D section.

6 Measurements and Characteristics of $Ta/Ta_2O_5/TiN$ ReRAM Cell

This chapter presents the measurement results and analog behaviour of tantalum/tantalum pentoxide/titanium nitrate ReRAM cell. Without proper understanding of the general characteristics and physical mechanism behind switching, it is impossible to tell if the device is suitable for the proposed task.

The purpose of this study is to analyse the analog behaviour of the ReRAM cell. Analog behaviour is achieved if multiple resistance states are obtained. Reliable analog operation is necessary if operation as artificial synapse is considered. The variable parameters are device area (diameter, ϕ) and oxide thickness, t_{ox}.

Measurements were carried out at Hokkaido University's Laboratory of Nanomaterial Science. The primary measurement device was Keysight's B1500A semiconductor device parameter analyser which provides good $i - v$ sweep and voltage pulse measurement features. Voltage is applied to Ta top electrode and TiN bottom electrode is a common ground for measurements. Devices with following hole sizes , $\phi = \{4\,\mu\mathrm{m}, 8\,\mu\mathrm{m}, 16\,\mu\mathrm{m}, 32\,\mu\mathrm{m}, 64\,\mu\mathrm{m}\}$ and oxide thicknesses, $t_{\mathrm{ox}} = \{7\,\mathrm{nm},\ 10\,\mathrm{nm},\ 20\,\mathrm{nm}\}$ were measured. Measurements were done in room temperature (22 °C) and under atmospheric pressure (0.1 MPa).

6.1 General Characteristics

A bipolar switching behaviour is observed in $Ta/Ta_2O_5/TiN$ cell where switching to LRS happens under positive voltage and switching to HRS happens under negative voltage. A simplified model and typical switching behaviour is shown in Figure 6.1. Initially after manufacturing process the device is in pristine state. There exists chemically formed oxygen deficient interface $Ta_2O_{5\text{-}\delta}$ between Ta and Ta_2O_5 layers where oxygen vacancies are present. Switching behaviour occurs when a conducting channel between Ta and TiN

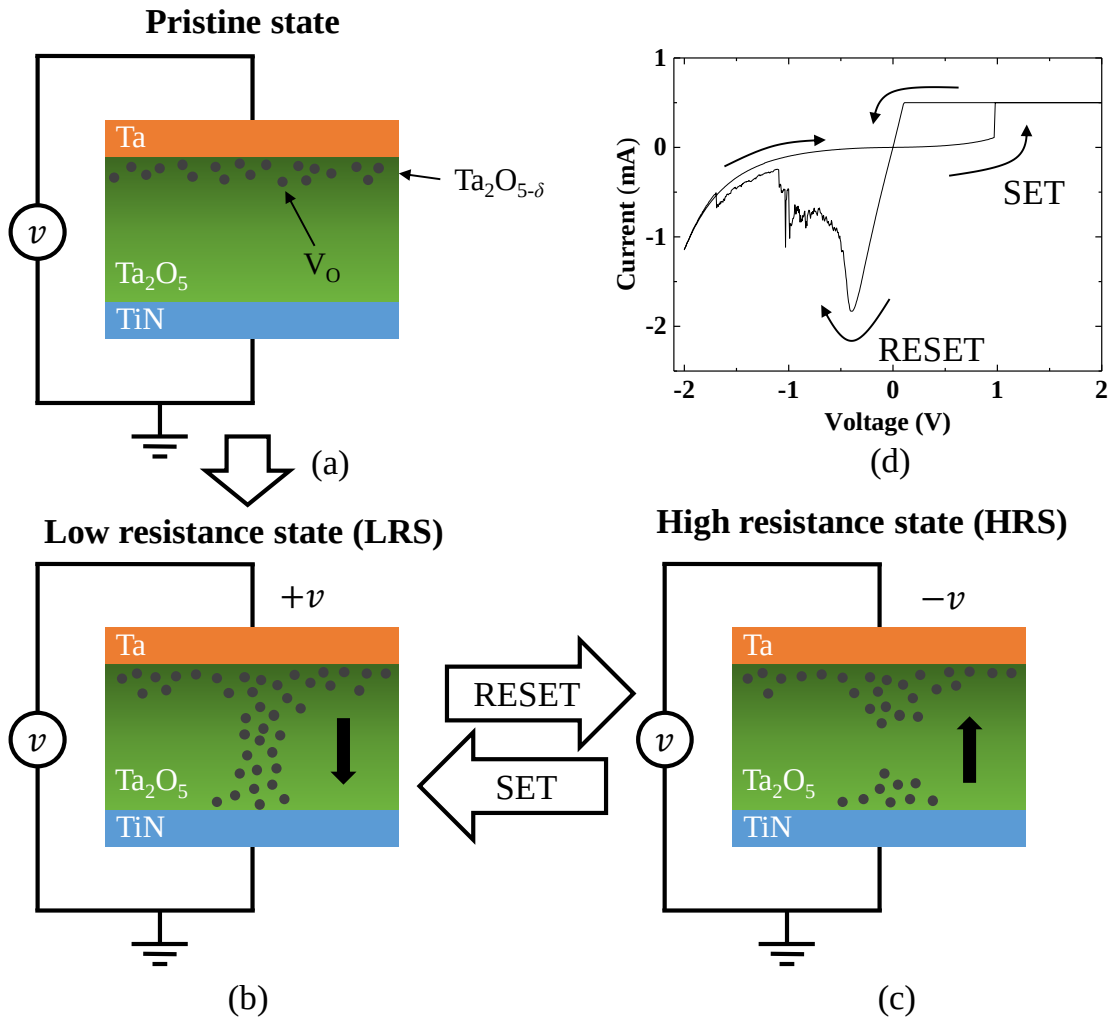

Figure 6.1: (a) Device in initial state, (b) device after forming and (c) device after reset. (d) Typical switching characteristics.

electrodes form. In this case, a redox-based reaction forms a conducting channel that consists of oxygen vacancies V_O. Basically this phenomenon can be said to be valence change memory effect (VCM) that is explained in Section 3.4.

Applying an electric field over Ta_2O_5 layer makes the oxygen vacancies to move and change the stoichiometry of the oxide material according to the reaction

$$2TaO_2 + O^{2-} \leftrightharpoons Ta_2O_5 + 2e^-.$$

When the electric field is strong enough, in other words when positive voltage crosses some threshold, an abrupt change in device's resistance is observed. This happens because oxygen vacancies are moved near TiN interface and form a fully conductive channel. ReRAM

cell is said to be in LRS.

Switching back to HRS requires an opposite electric field which means that voltage with opposite polarity must be applied. Small negative voltages are not enough to change the device's state but at some point the electric field becomes dominant and oxygen vacancies begin to move back towards Ta electrode. A gradual change in resistance is observed and when the conductive channel is fully broken the device returns back to HRS. However, the oxygen deficient layer is never as uniform as it was in pristine state. In following switching cycles the forming voltage should be lower.

6.1.1 Initial Forming Voltage and Resistance

It is expected that initial forming voltages and initial resistance values are higher than in subsequent switching cycles. This effect is studied by analysing the experimental data gathered from devices that has different t_{ox} and ϕ. Figure 6.2 shows that there exists a clear linear dependence between t_{ox} and initial forming voltage as well as initial resistance. All the devices that have different area are gathered in these two graphs to show also the statistical variation. It is interesting to see that the initial forming voltage is already $< 1.0\,\mathrm{V}$ when $t_{ox} = 7\,\mathrm{nm}$. This implies the possibility of forming-free devices when $t_{ox} \leq 5\,\mathrm{nm}$. However, this does not come without a trade-off. As it can be seen on the same figure the initial resistance becomes very low when oxide thickness is reduced. In the case of $t_{ox} = 7\,\mathrm{nm}$, the median initial resistance value is in the range of $1\,\mathrm{k}\Omega$ although there is about two orders of magnitude variance due to device area. The initial resistance is obtained by measuring the current at $10\,\mathrm{mV}$ and calculating the resistance according to Ohm's law. This method should give an accurate value because in the low voltage region the $i - v$ characteristics are assumed to be linear. Figure 6.2 (b) also shows that B1500A cannot measure the current accurately when resistance values are larger than $10^{10}\,\Omega$.

The area of device has also an impact on the forming voltage and resistance values as shown in Figure 6.3. Forming voltage changes slightly when area is increased while t_{ox} has much greater effect. Resistance values are clearly inversely proportional to the area because initially no filament exists and current flows through the entire region of the insulator. Figure 6.3 also shows why nanoscale devices are necessary when t_{ox} is decreased. Experimental values demonstrates that initial resistance is too low ($\approx 100\,\Omega$) when $t_{ox} = 7\,\mathrm{nm}$ and device area is larger than $1\,000\,\mu\mathrm{m}^2$. It is the reason why experimental measurements of $\phi = 64\,\mu\mathrm{m}$ cells are missing because low initial resistance limits greatly

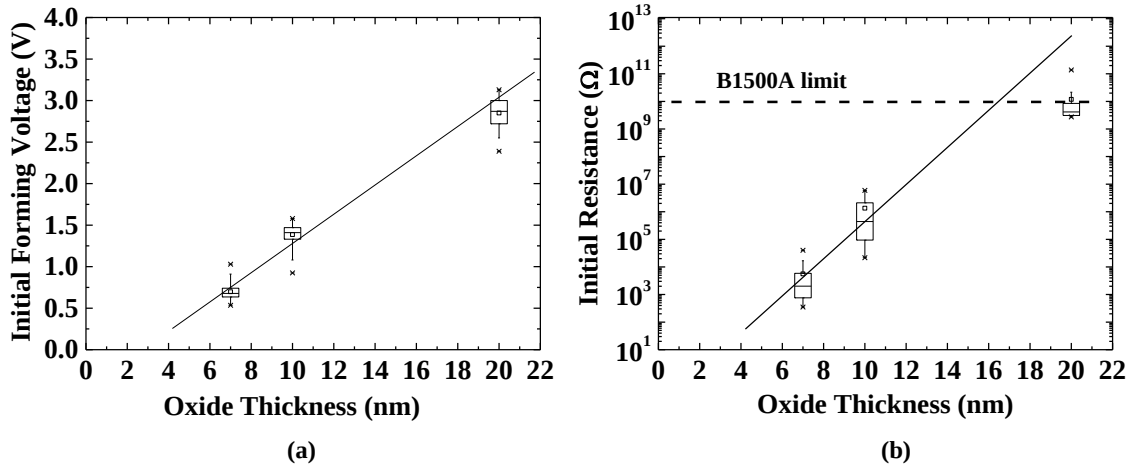

Figure 6.2: (a) Linear dependence between t_{ox} and initial forming voltage. (b) Linear dependence between t_{ox} and initial resistance.

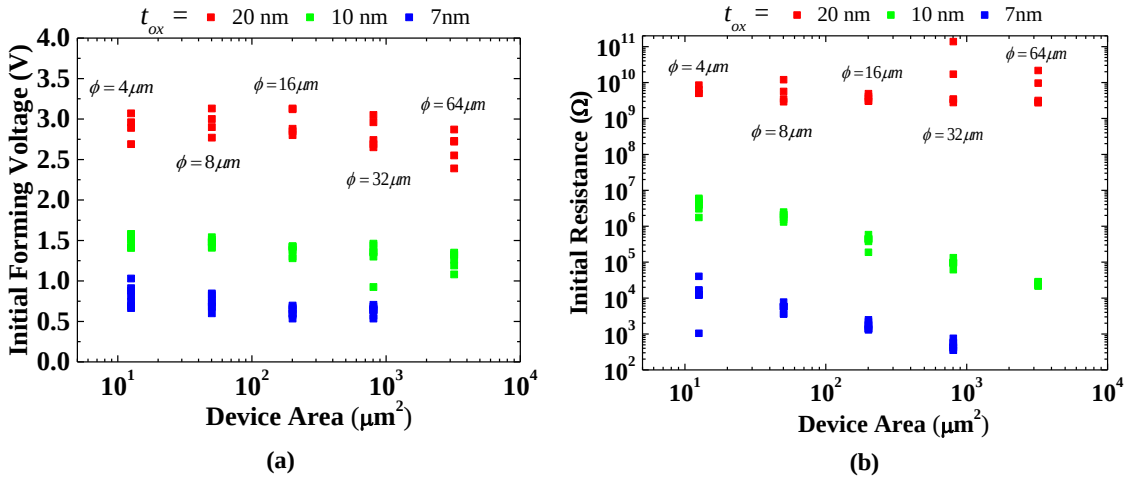

Figure 6.3: (a) Area vs. initial forming voltage. (b) Area vs. initial resistance. A corresponding ϕ of the area is also shown.

$R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ and no switching behaviour is observed. Thinner oxide layers probably work only with devices whose dimensions are in scale of nanometers.

This behaviour can be explained if the simplified model in Figure 6.1 is analysed. The oxygen deficient $Ta_2O_{5-\delta}$ layer is formed by chemical interactions thus the high density of oxygen vacancies fill the whole interface between Ta and Ta_2O_5. The area of the interface is assumed to be equal to the device's area. If it is also assumed that the thickness of $Ta_2O_{5-\delta}$ layer is the same regardless of t_{ox}, smaller electric field is necessary to form a conducting channel between electrodes in thinner oxide materials.

The geometry and properties of conducting filament as well as physical switching mechanisms of proposed ReRAM cell are discussed more deeply in Section 6.2.

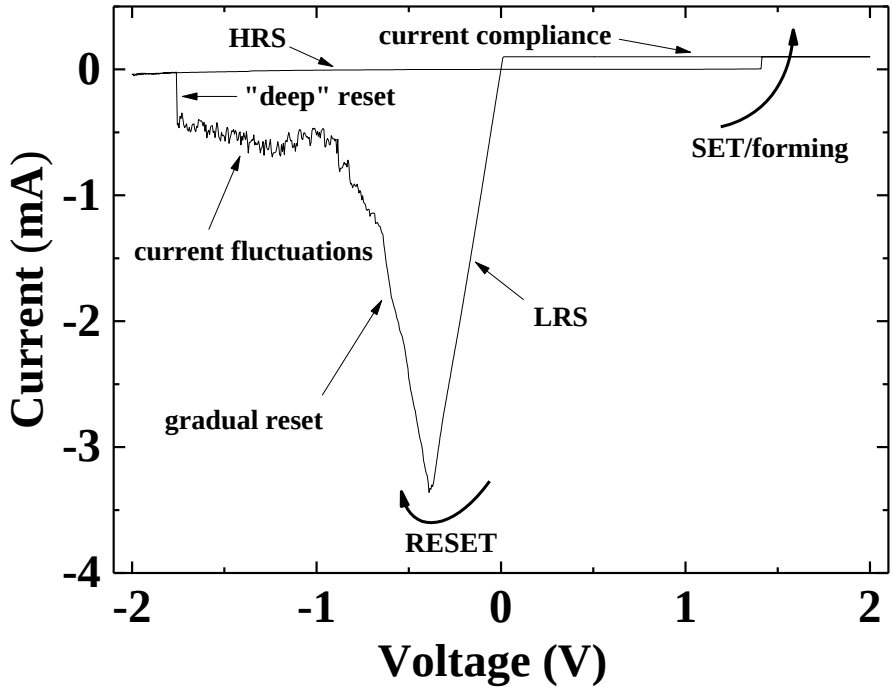

Figure 6.4: Typical $i - v$ characteristics of Ta_2O_5-based ReRAM cell.

6.1.2 Typical Current-Voltage Characteristics

The $i - v$ characteristic gives a good overview how ReRAM devices behave. Typically, $i - v$ curves are obtained by DC voltage sweeps, that can be interpret as triangle shaped low frequency voltage pulses. Proper current compliance is set with B1500A. Similar type $i - v$ characteristics are observed regardless of device area or oxide thickness although some differences exists.

In Figure 6.4, the $i-v$ curve of ReRAM device with $t_{\text{ox}} = 10\,\text{nm}$ and $\phi = 4\,\mu\text{m}$ is shown. Initially the device is in HRS. When positive voltage sweep is performed an abrupt change to LRS happens above 1.0 V and current compliance is necessary to prevent breakdown. Reset process happens during the negative voltage sweep. Current increases linearly at low negative voltages. Reset occurs around $-0.5\,\text{V}$ when current starts to decrease sharply. Between $-0.5\,\text{V}$ and $-1.0\,\text{V}$ a gradual increase in resistance is observed before the region of random current fluctuations. A sharp current drop typically happens between $-1.5\,\text{V}$ and $-2.0\,\text{V}$. This can be interpret as complete filament cut-off between electrodes and thus it is called "deep" reset or full reset. The device returns to HRS. Although this example is called typical $i - v$ curve the exact current and voltage values of different regions and operating points vary greatly even in one device during subsequent sweeps.

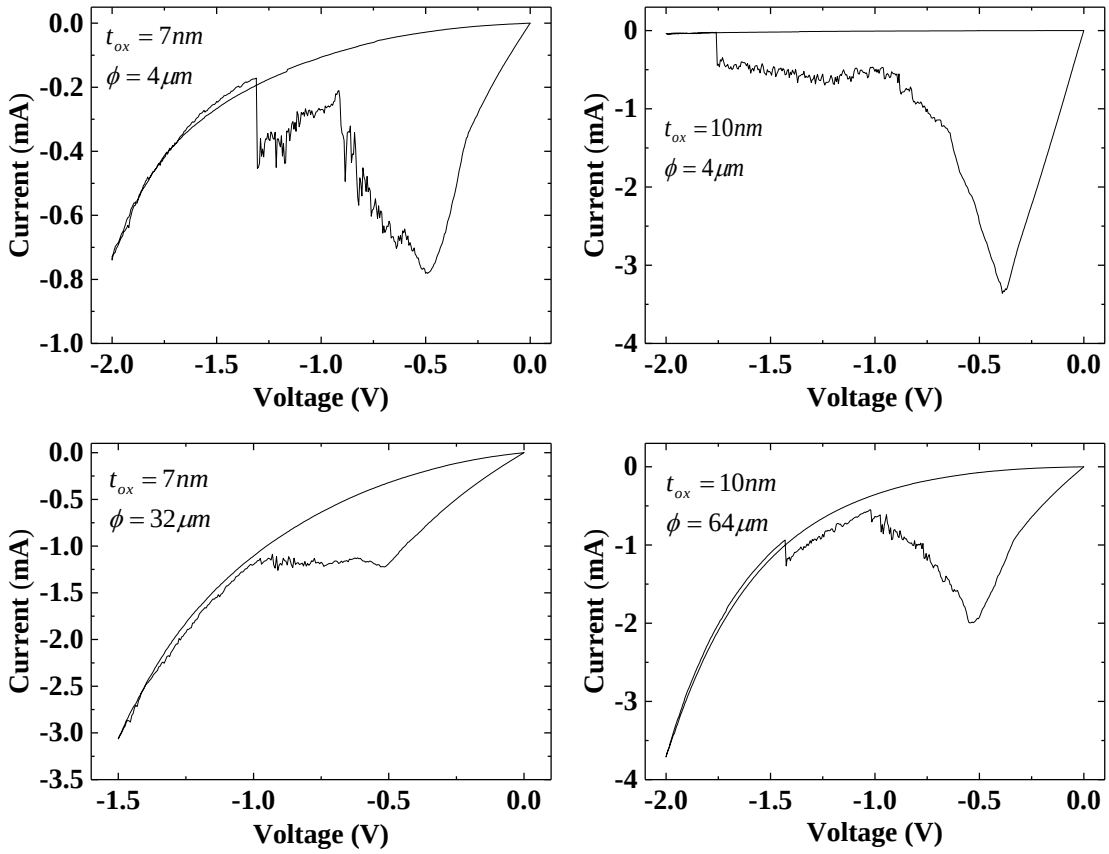

Figure 6.5: Four $i - v$ curves of devices with different dimension.

Different oxide thicknesses and device areas have different type of $i - v$ curves. Four different examples of reset behaviour are shown in Figure 6.5. Decreasing t_{ox} and increasing device area has an exponential effect on HRS current behaviour. Exponential current increase is undesirable effect as it increases the power consumption significantly. It also limits $R_{\text{OFF}}/R_{\text{ON}}$ because the exponential behaviour is already observed at low voltages especially when $t_{\text{ox}} = 7\,\text{nm}$. Metal-insulator-metal stack has always some leakage current going through. The structure is equivalent to a capacitor with a capacitance $C = \epsilon_r \epsilon_0 \frac{A}{d}$, where ϵ_r is the relative permittivity of the oxide material, ϵ_0 is the vacuum permittivity, A is the capacitor area, and d is the distance between electrodes. In the case of the ReRAM cell, $A = \frac{1}{4}\pi\phi^2$ and $d = t_{\text{ox}}$. The relative permittivity is rather high for Ta_2O_5 ($\epsilon_r \backsim 25$) compared to SiO_2 ($\epsilon_r \backsim 3.9$). High ϵ_r increases ReRAM capacitance significantly but also reduces the leakage current efficiently. Capacitance affects ReRAM's frequency response because high capacitance attenuates high frequency voltage pulses. The leakage current mechanism is supposed to be similar to MOSFET gate leakage current where current flows through gate oxide (typically SiO_2). Dominant mechanisms are analysed in Section 6.2.

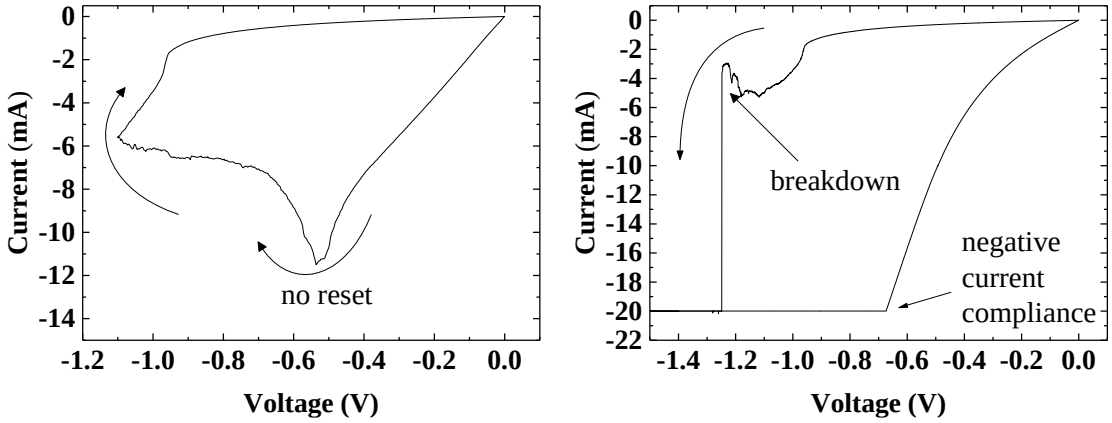

Figure 6.6: Reset behaviour of $t_{ox} = 20\,\text{nm}$ devices.

The reason why all $t_{ox} = 20\,\text{nm}$ devices are missing from this examination is that no resistive switching cannot be observed regardless of cell size. All devices are able to switch from initial HRS to LRS but switching from LRS to HRS does not happen before breakdown. Breakdown could be prevented by stopping the negative voltage sweep at low voltages. It seems that a partial reset process occurs but on the subsequent positive voltage sweep no forming is observed which leads to a conclusion that higher negative voltages are needed for a complete reset. It turns out to be impossible as typical breakdown voltage is between $-1.1\,\text{V}$ to $-1.4\,\text{V}$. Examples of reset behaviour are presented in Figure 6.6. Breakdown is observed when current sharply drops before it is limited by negative current compliance. Broken devices have short circuit like behaviour from which they can not recover.

In this study the only useful information about $t_{ox} = 20\,\text{nm}$ devices are initial R_{OFF} and R_{ON} values as well as initial forming voltages.

6.1.3 Understanding the Breakdown Mechanisms

Two types of different breakdown mechanisms are observed during $i - v$ sweep. One is the previously mentioned reset breakdown when abrupt increase of negative current occurs suddenly. Another type of breakdown happens during positive voltage sweep during forming. The forming process is so fast that the B1500A current limitation circuit does not have time to prevent current overshoot. The magnitude of overshoot current is not known but obviously it is high enough to cause a permanent breakdown. The following reset sweep after overshoot does not show any kind of switching behaviour as negative

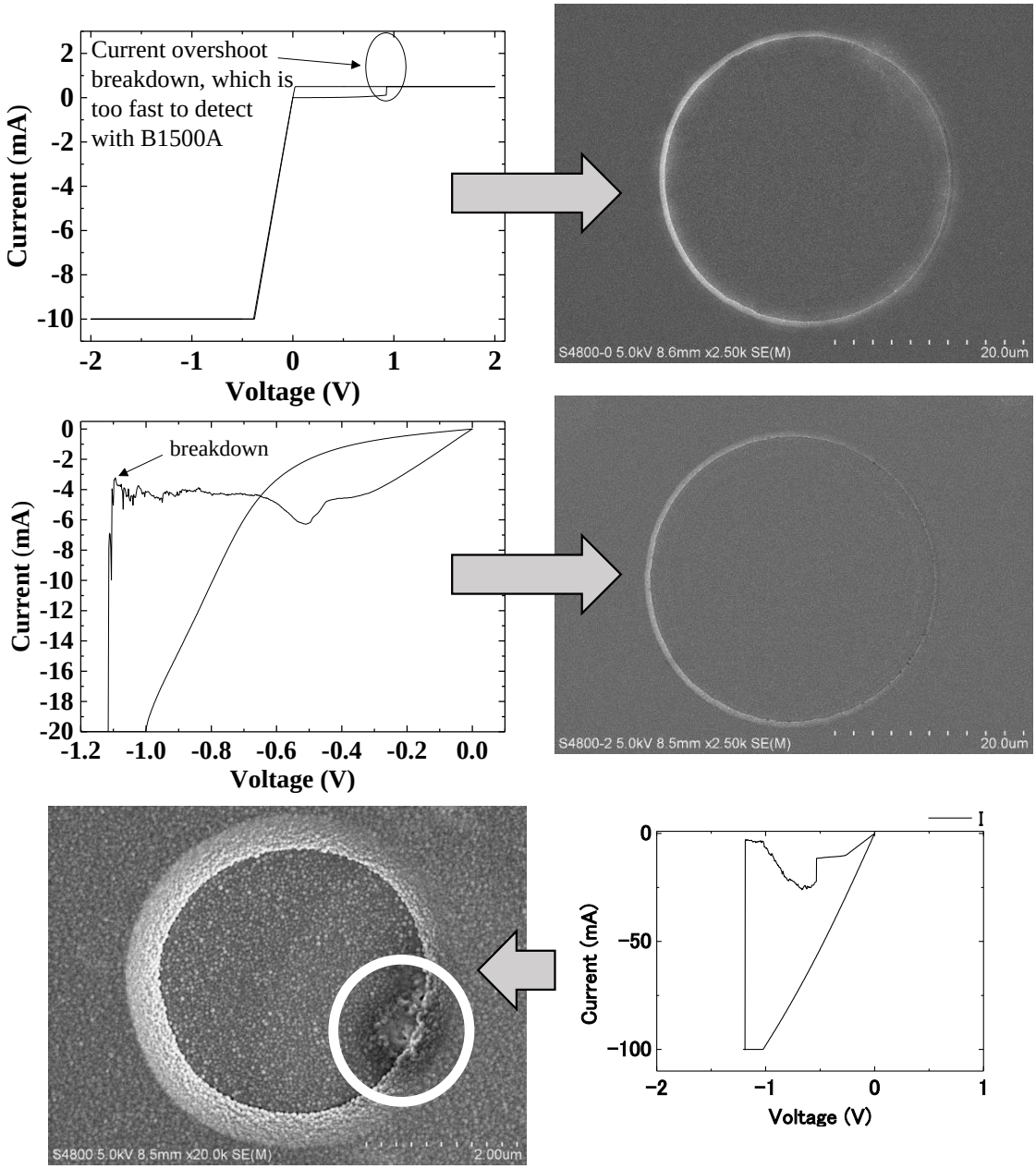

Figure 6.7: SEM images of broken devices. The top two images do not show any damages as SEM images of working devices were similar. The lowest SEM image is taken from $Cu/Ta_2O_5/TiN$ device for comparison. A clear evidence of reset breakdown can be seen.

current increases steadily before it is limited by compliance current.

Some of the broken devices were studied using *scanning electron microscope* (SEM) to see if there is any clear physical evidence of breakdown. Figure 6.7 shows SEM images of broken devices where two different types of breakdown mechanisms occurred. Any visible damage can not be seen. In the same figure, another device made of $Cu/Ta_2O_5/TiN$ is shown for comparison. A clear area where breakdown happened is located near the edge

of the device. Although the compliance current level was set much higher the probable reason for visible damage may be due to different filament geometry. In theory a switching mechanism of device that has electrochemically active metal, in this case Cu top electrode, is based on electrochemical metallization effect (ECM). The conducting filament is made from Cu^+ ions instead of oxygen vacancies thus the current densities as well as filament geometry is different.

Even if the current overshoot is not large enough to cause permanent breakdown it is possible that still some damages occur inside the oxide material or electrodes. Minor damages have an effect on cell's overall performance and over time progressive degeneration leads to unreliable operation. A strict current control seems to be mandatory for the manufactured device in this study. A MOSFET current control and 1T1R cell was manufactured and results are presented in Chapter 7.

6.1.4 Relationship Between Forming and Reset Processes

In theory, the forming process should define the properties of filament. A conducting filament has a "strength" that needs to be broken during reset process. Figure 6.8 (a) shows the two points in the $i-v$ curve where experimental results from multiple devices are gathered. Forming current is measured at the point where forming occurs and reset current is the maximum current before reset process happens. In the case of $Ta/Ta_2O_5/TiN$ cell, reset current seems to be relatively high as it is typically more than 1 mA. Observations of multiple $i-v$ curves lead to a conclusion that there was a clear dependence between forming voltage and reset current when $t_{ox} = 10\,nm$ as Figure 6.8 (b) suggests. Higher forming voltages lead to higher reset current linearly. Device area does not have any noticeable effect on the linear relationship between forming voltage and reset current as scattered data points suggest. In the case of $t_{ox} = 7\,nm$, no correlation between forming voltage and reset current exists. The reason behind this is currently unknown.

Another interesting observation is the relationship between forming voltage and the corresponding current at the point where abrupt forming happens. Forming current decides the limit for minimum current compliance level. If lower power operation is desired this limit should be as low as possible, preferably below 1 mA.

Figure 6.8 (c) shows all measured forming currents as a function of forming voltage. Oxide thickness and device size are clearly distinguishable from the graph. The exponential behaviour of HRS current, that is previously explained and shown in Figure 6.5,

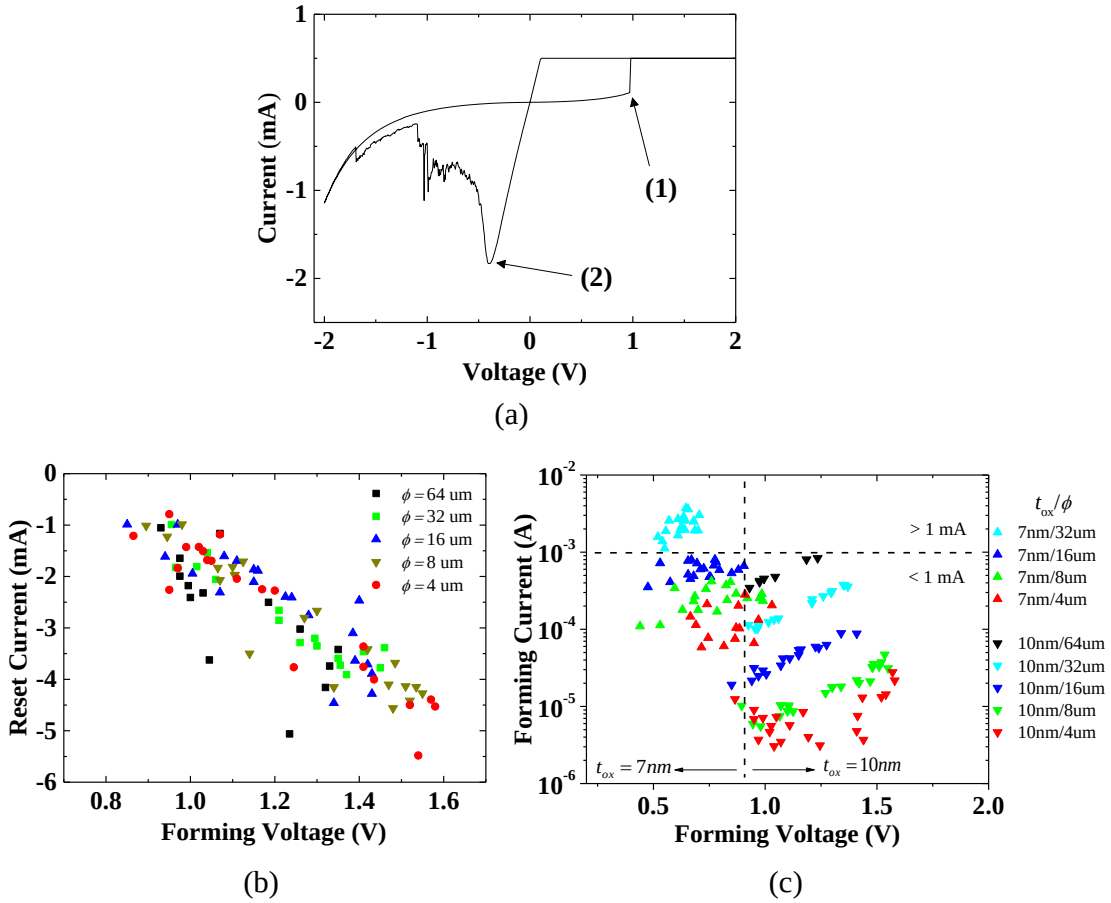

Figure 6.8: (a) Two operating points in $i-v$ curve that define: (1) forming voltage and forming current, (2) reset voltage and reset current. (b) Linear dependence of forming voltage vs. reset current when $t_{\mathrm{ox}} = 10\,\mathrm{nm}$. (c) Relationship between forming voltage vs. forming current.

affects to the point where forming happens. Lower R_{OFF} in HRS leads lower forming voltages but also higher forming currents that limits the minimum current compliance levels. Decreasing oxide thickness and increasing device area will result higher forming currents and especially in the case of $t_{\mathrm{ox}} = 7\,\mathrm{nm}$ any device that has $\phi \geq 32\,\mu\mathrm{m}$ needs compliance current over 1 mA. If thin metal oxide layers are needed the only option is to reduce device area.

The negative voltage level where reset process begins is important if analog operation via reset voltage modulation is considered. Ideally reset process should be identical between every subsequent cycle which makes the operation of ReRAM cell predictable. Studying the $i-v$ curves shows that in this case the reset voltages are very similar regardless of oxide thickness or area. It also shows that higher forming voltage or current does not have any remarkable effect on the reset voltage which is somehow surprising result as it is natural to assume that stronger electric fields are needed to break the filament.

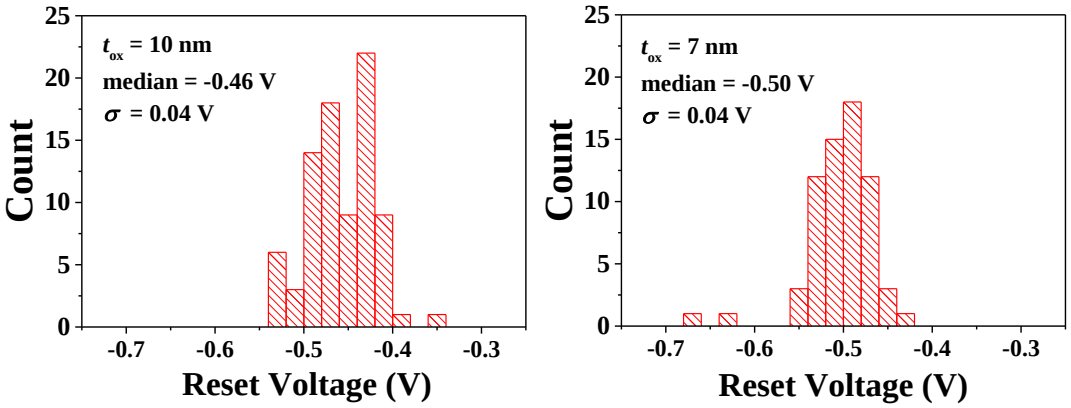

Figure 6.9: Gaussian distributions of reset voltages. Corresponding t_{ox}, median and standard deviation values are shown.

Figure 6.9 shows the histogram graphs of reset voltages. Measurements of multiple reset voltages seem to follow Gaussian distribution.

6.1.5 Retention

Long data retention time of ReRAM cell is one of the requirements that is mentioned in Section 4.3. However, there was a problem, probably due to measurement device that caused multiple breakdowns during measurements. The measurement arrangement was to set multiple devices into LRS and measure R_{ON} at different times. This is the reason why retention measurements are not analysed in this thesis. Only one retention time of single device was obtained experimentally. This device kept it's R_{ON} for 10^6 s (last measurement point) which is still far less than the requirement for non-volatile memory to store its data for ten years ($\backsim 10^8$ s).

6.1.6 Endurance and $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$

At least 10^3 endurance cycles are needed to compete with certain applications of Flash technology. Over 10^{16} cycles is a requirement for DRAM but not a single study has been able to produce such values yet. The main focus of this study was to analyse analog behaviour of ReRAM cell. For this reason the endurance cycles are too few to be comparable with other studies or Flash. Endurance is mainly studied to understand the general characteristics such as $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$.

Endurance, R_{OFF} and R_{ON} values are obtained from the $i-v$ sweep by measuring the current at $\pm$ 10 mV and calculating the resistance values according to Ohm's law. Devices that work after 100 set and reset sweeps are considered well behaving. If there

is no degeneration in R_{OFF} and R_{ON} values data can be extrapolated up to 1 000 cycles. However it is not uncommon that devices break down suddenly without any degeneration so extrapolation may not reveal the real performance. Another way to measure endurance value is to apply voltage pulses which is more desirable approach because it provides faster operating speed and energy efficiency. For practical applications, pulse programming is the only option as DC sweeps like the ones that measurement systems provide are too slow thus increasing the energy consumption considerably. The main reason that voltage pulse measurements are missing from this review is explained in Section 6.1.7.

Measurement data showed that many devices are able to survive 100 set and reset sweeps. No degeneration is observed during 100 sweeps and there are two well distinguishable resistance states. During measurements some devices were broken mainly due to overshoot current that is explained in breakdown mechanisms section. Figure 6.10 shows rather typical behaviour of two different devices. In the case of $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ devices the absolute resistance values as well as variance of resistance states is higher than $t_{\mathrm{ox}} = 7\,\mathrm{nm}$ devices. However $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ has greater value. There seems to be a trade-off between resistance variance and $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ which is confirmed in statistical analysis of multiple endurance measurements.

Analysing the statistical variance where results of all measured cells are gathered together reveals some properties that are necessary to considered in analog operation. Figure 6.11 shows the distribution of R_{OFF} and R_{ON} and also a rough approximation of $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ value. All the $t_{\mathrm{ox}} = 7\,\mathrm{nm}$ devices, expect $\phi = 4\,\mu\mathrm{m}$, have very little variation on their resistance values. Controllability and repeatability is considered an important factor for ReRAM cells. Thinner oxide has also the relatively low $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$. If lower power operation is considered higher R_{OFF} values as well as low leakage in HRS state during read operations is better in $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ devices. R_{ON} seems to be independent of the oxide thickness as the magnitude is almost the same.

It seems that t_{ox} has a big impact on $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$. It is hard to interpret the results of $t_{\mathrm{ox}} = 20\,\mathrm{nm}$ because they lack resistive switching property but still the initial R_{OFF} and R_{ON} can be calculated from the first set sweep. The calculated value $R_{\mathrm{OFF}}/R_{\mathrm{ON}} \approx 10^6$ shows that t_{ox} has an exponential factor as the ratio is many orders of magnitudes higher than in $t_{\mathrm{ox}} = 7\,\mathrm{nm}$ and $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ devices. In theory, greater $R_{\mathrm{OFF}}/R_{\mathrm{ON}}$ should be better if analog operation is considered as it provides bigger operating window for intermediate resistance states. In practice this may not be the case because variance between resistance

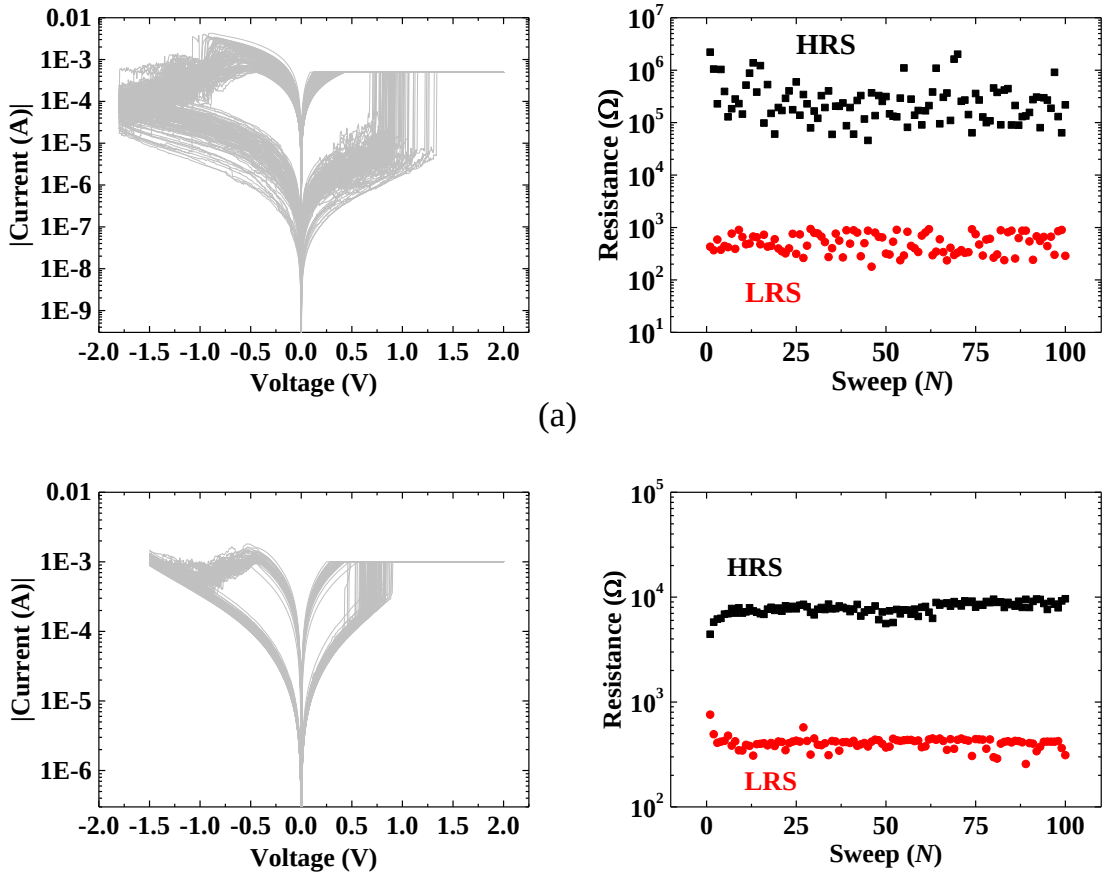

Figure 6.10: $i - v$ curves of 100 set/reset sweeps as well as corresponding resistance values. Two resistances states (HRS and LRS) are clearly distinguishable. Device dimensions are: (a) $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ and $\phi = 4\,\mu\mathrm{m}$, (b) $t_{\mathrm{ox}} = 7\,\mathrm{nm}$ and $\phi = 8\,\mu\mathrm{m}$.

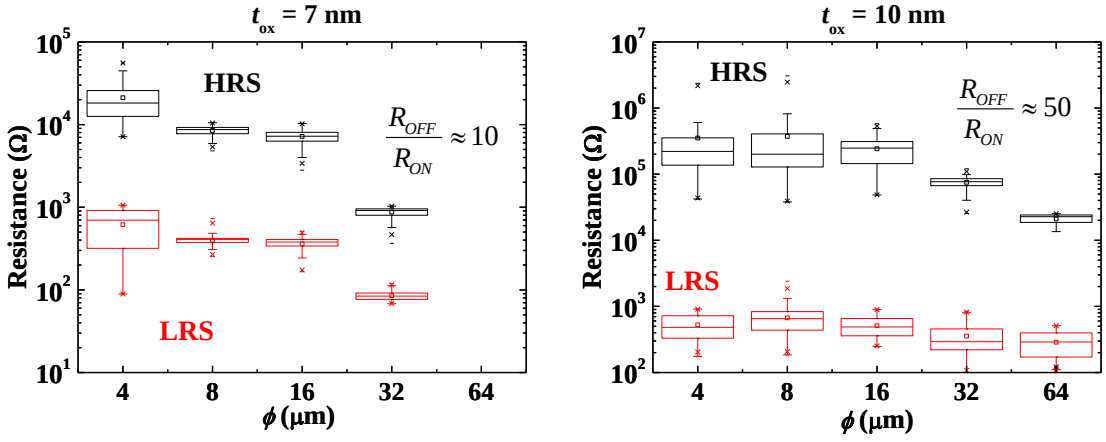

Figure 6.11: Statistical distribution of R_{OFF} and R_{ON} values.

values grow. This seems to affect Ta_2O_5 ReRAM cells. An interesting choice for further studies would be to investigate intermediate oxide thicknesses such as 15 nm to confirm the link between resistance variability, R_{OFF}/R_{ON} as well as examining if such t_{ox} values are able to carry out switching.

6.1.7 Pulse Measurements

For pulse measurements, a waveform generator/fast measurement unit (WGFMU) was used with B1500A. The unit has measurement current limit to perform accurate current measurements at the expense of maximum current level. The unit does not provide any current compliance. The lack of proper current compliance setting was the main reason of multiple breakdowns during measurements. The higher voltage pulse amplitudes resulted current overshoot breakdowns and experimental results are very limited to interpret any meaningful data from pulse measurements.

A positive voltage pulse that has an amplitude over 1.0 V is typically enough to perform a switching from HRS to LRS. However the change is so abrupt that even the current overshoot is too fast to detect with measurement unit that has the sampling rate of 5 ns. The fast on-switching time (< 10 ns) is consistent with literature [61].

Figure 6.12 shows two different current measurements of reset pulses. In Figure 6.12 (a), a reset voltage pulse whose rise time, fall time and pulse width are 20 µs and amplitude is -1.8 V shows the reset operation in time domain. The regions that are presented also in Figure 6.4 can be clearly seen. Reset process begins around -0.5 V and current drops sharply until it reaches the current fluctuation region. A "deep" reset happens around -1.8 V at the same time when the reset pulse reaches its peak. There are minor current fluctuations in HRS before voltage is dropped back to 0 V. Interesting observation here is the off-switching time which is about 20 µs that is many times longer than on-switching time. This behaviour is also expected and reported in literature [61].

Figure 6.12 (b) shows the problem in pulse measurement. This time reset process is stopped before it fully happens. -0.6 V pulse is applied many times using fast rise and fall times. Clear current spikes are visible in the graph. If these current spikes cannot be removed, devices are prone to premature breakdowns. Without proper current compliance circuit pulse measurements are impossible to perform. As it was previously mentioned, pulse operation is preferred over DC sweep due to operation speed and energy efficiency. This is also the reason why a MOSFET is studied in Chapter 7.

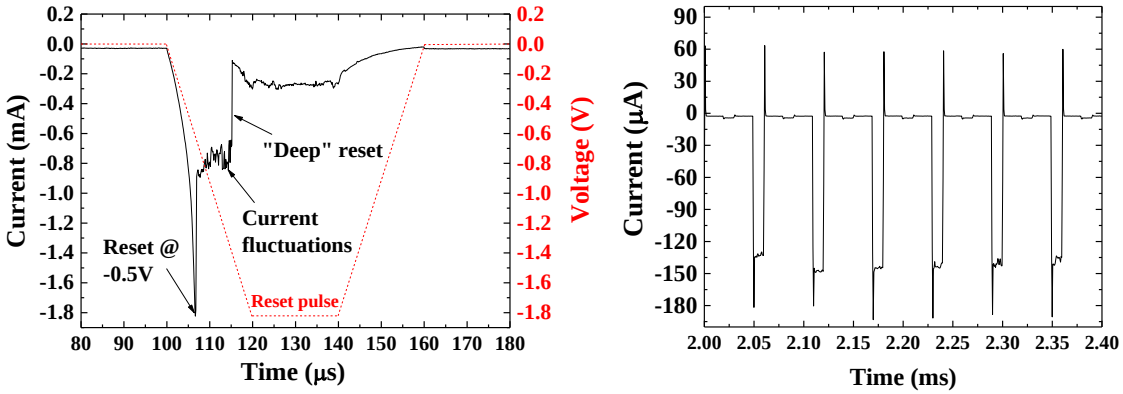

Figure 6.12: (a) Typical behaviour of one reset pulse. (b) Multiple reset pulses where clear current spikes are visible. Reset pulse amplitude is $-0.6\,\mathrm{V}$.

6.2 Physical Mechanisms Behind Switching

The conduction mechanisms in ReRAMs are already discussed in general level in Chapter 3. In this section, the approach is to analyse the conduction mechanisms in the considered $Ta/Ta_2O_5/TiN$ structure. Conductions mechanism can be different in HRS and LRS, but in the following HRS is considered more carefully since in this study the analog operation relies on the modulation of the reset voltages. It is not uncommon to observe different conduction mechanisms and $i-v$ characteristics even in the same materials. Many different attributes have an effect e.g. the pureness of electrode and oxide materials, annealing temperature in fabrication process and operating temperature or other ambient conditions. One classification of conduction mechanisms in dielectric films is presented in [74], where they are divided into *electrode limited* and *bulk limited* mechanisms. Mechanisms among the electrode limited are Schottky emission, Fowler-Nordheim tunneling, direct tunneling and thermoionic-field emission. Among bulk limited are Poole-Frenkel emission, hopping conduction, ohmic conduction, space charge limited conduction, ionic conduction and grain-boundary-limited conduction.

If reliable analog operation is considered, repeatable reset behaviour would be the most important criteria in this study. In the case of manufactured $Ta/Ta_2O_5/TiN$ structure there could be multiple mechanism that affect $i-v$ characteristics but typically one mechanism is more dominant than others. Especially the current fluctuations and leakage current in HRS are analysed to understand the behaviour. If a dominant mechanism and a suitable physical model is found there is also a possibility to construct a simulator model

of the device to be used in *computer-aided design* (CAD) tools.

Analysing the $i - v$ characteristics, especially during the reset process, reveal some dominant conduction mechanism. Observations are very well consistent with *space charge limited conduction* (SCLC). The fingerprint of SCLC is three regions of distinct slopes in the $i - v$ curve when plotted in log-log scale. These regions are (i) *Ohmic region* ($i \propto v$), (ii) trap-unfilled region ($i \propto v^2$) and (iii) trap-filled region also known as *Child's square law region* ($i \propto v^2$). Ohmic region is observed at low electric fields while the square law region is dominant at higher electric fields. The current density equations are [74],

$$\text{(i)} \quad J = en_0\mu\frac{v}{d}, \tag{6.1}$$

$$\text{(ii)} \quad J = \frac{9}{8}\mu\epsilon_r\epsilon_0\theta\frac{v^2}{d^3}, \tag{6.2}$$

$$\text{(iii)} \quad J = \frac{9}{8}\mu\epsilon_r\epsilon_0\frac{v^2}{d^3}, \tag{6.3}$$

where n_0 is the concentration of the free charge carriers in thermal equilibrium, μ is electron mobility and θ is the ratio of the free carrier density to total carrier density. Rest of the symbols are as previously defined. In ReRAMs, Child's law region (6.3), is observed if current abruptly increases in HRS but in this case the measured devices follow (6.2) which is the area of strong injection when the traps are not totally filled yet. Ohmic region (6.1) is typical at low voltages when $|v| < 0.1\,\text{V}$.

Figure 6.13 shows typical $i - v$ behaviour of two chosen devices in log-log scale. Exponential current increase is obvious in HRS when reset voltage is higher than $-0.1\,\text{V}$. Ohmic region ($i \propto v$) is also present in both LRS and HRS. Sometimes in LRS, the current starts to increase exponentially just before reset happens. This observation is also very well consistent with the conclusion that the most dominant conduction mechanism in this study is indeed SCLC. Analysing the mechanisms during positive voltage sweep are expected to be the same even though compliance current hides the behaviour in LRS.

It is not clear if the filament is totally cut-off between electrodes in full reset. Existence of very narrow filament in Ta_2O_5 could be justified by the scattered R_{OFF} values in HRS which is observed in many devices especially when $t_{\text{ox}} = 10\,\text{nm}$. It is intuitive to presume that R_{OFF} should be nearly the same between subsequent reset cycles if filament is completely cut-off.

Similar $i - v$ characteristics compared to this study has been reported in literature for

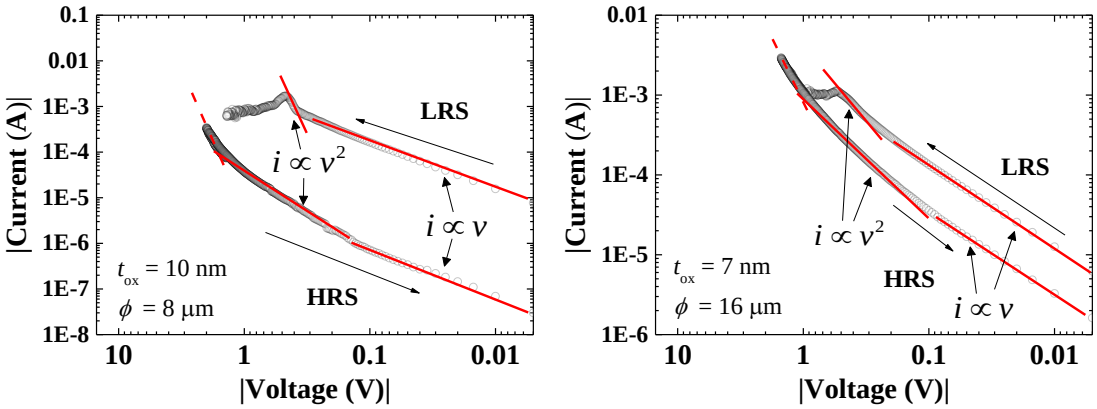

Figure 6.13: Log-log curves of reset sweeps. $i - v$ characteristics shows typical SCLC behaviour. Ohmic region ($i \propto v$) and trap-unfilled region ($i \propto v^2$) are visible.

TaO_x-based [43, 75], and also for HfO_x-based [76], ReRAM cells. Those studies used a *quantum point contact* (QPC) model to explain the device $i - v$ characteristics especially in HRS. QPC model was originally proposed to model the conduction after soft or hard breakdowns in MOSFET oxides. QPC model assumes that there exists a filament in oxide and its narrowest point the first quantized sub-band acts as a potential barrier for charge carriers. A complete theory and its application for ReRAM is explained by Miranda et al. in [77], and [78], respectively. The current in QPC model can be approximated to be

$$i \approx \frac{2e}{h} N \left(ev + \frac{1}{\alpha} \ln \left(\frac{1 + \exp(\alpha\Phi - \alpha ev/2)}{1 + \exp(\alpha\Phi + \alpha ev/2)} \right) \right), \tag{6.4}$$

where e is the elementary charge, h is the Planck's constant, N is the number of conducting channels, Φ is the potential barrier height measured with respect to Fermi level E_f, and α is the shape parameter related to parabolic shape in the energy band diagram at E_f. Approximation (6.4) is suitable for modelling HRS ($N = 1$) and LRS ($N > 1$) current and it is used successfully for modelling HRS current in HfO_x at low voltages $|v| < 0.4\,\mathrm{V}$ in [79], and showed excellent fitting in wide voltage and resistance range in [80]. Slightly modified model is also used for TaO_x-based devices in [43, 75].

Fitting QPC model by modifying Φ and α parameters to approximate experimental values obtained in this study is somehow problematic. Model is able to explain the behaviour only at low voltages, $|v| < 0.1\,\mathrm{V}$. Especially good fitting that is shown in [80], is revealed in log-log $i - v$ curve where current increases linearly up to 1 V. In this study the measured devices show log-log linear current behaviour usually up to 0.1 V in HRS. At higher voltages the QPC model is not very accurate and it may not be suitable for

modelling the devices in this study.

Exponential leakage current can be explained with SCLC theory but the gradual reset slope before current fluctuations should be analysed as well. Studying the $i - v$ curves shows that sometimes current drops in steps under high electric fields before "deep" reset happens. Discrete resistance levels that are observed at the progressive reset phase are also reported in other studies [80,81], and those steps are interpreted as *conductance quantization* phenomena. A good review of background theory and observations in ReRAMs is given in [82]. The quantized unit of electrical conductance is defined as

$$G_0 = \frac{2e^2}{h} \approx 7.748091734625 \cdot 10^{-5}\ \mathrm{S}, \tag{6.5}$$

or in other words $G_0 = (12.9\,\mathrm{k\Omega})^{-1}$. Because the current fluctuations are very large, some statistical analysis is usually needed to reveal discrete conductance levels nG_0, where $n = 1, 2, 3, \ldots$. However, quantized conduction effect is only visible if the narrowest point in conductive filament is small enough. Considering pseudo-analog ReRAM operation the quantized conduction has its pros and cons. Discrete conductance levels can be produced more reliably between subsequent cycles if filament geometry is well defined. On the other hand, discrete levels are very limited because conductance nG_0 grows very large for relatively small n if conductances near HRS are considered.

Conductance quantization is studied to see if measured cells are affected by this particular phenomena. In Figure 6.14 some clearly visible discrete conductance levels are present in progressive reset phase. The exponential current behaviour is very dominant at lower end of conductance values which makes it difficult to interpret results. To confirm the theory, one hundred reset sweeps are analysed to see if conduction quantization affects the reset process. Results are shown in Figure 6.15. If conduction quantization phenomena is present during reset sweep they should be at least couple of discrete normalized conductance levels should be distinguishable from current fluctuations. Figure 6.15 (a) shows only randomly behaving current fluctuations. If such levels exists they should be revealed in statistical analysis. Histogram in Figure 6.15 (b) that there are no clear spikes in integer multiplies of G_0 that should follow Gaussian distribution. Although discrete conductance levels are clearly present in many devices the conductance quantization is not the cause. Any other explanation of such behaviour is not found.

The purpose of this section is to have a better understanding of the physical mech-

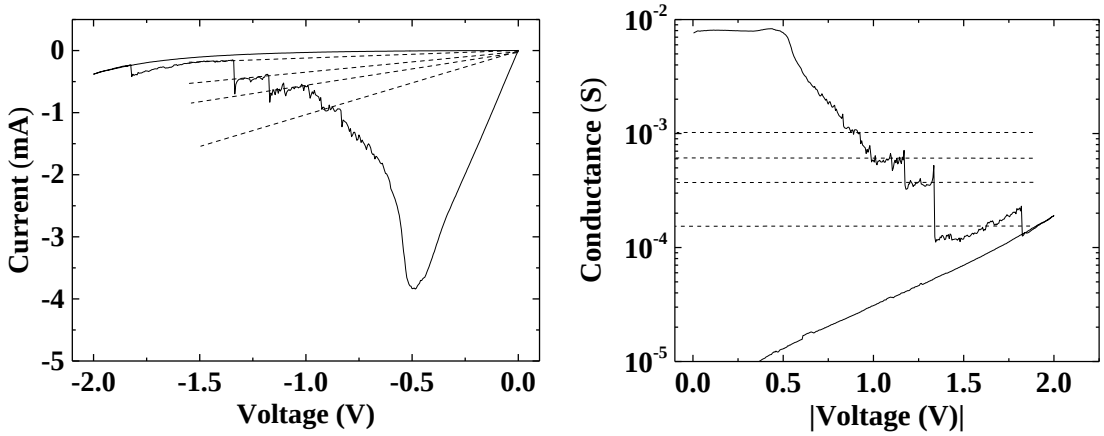

Figure 6.14: Discrete conductance levels are observed in these two graphs. Graph on the left plots the calculated conductance as a function of voltage. ReRAM cell dimensions are $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ and $\phi = 4\,\mu\mathrm{m}$.

anisms behind switching, especially during reset process that is main method to achieve multiple resistance states. Even though the QPC model and conductance quantization explain similar type of $i - v$ characteristics in other studies e.g. [42, 43, 75, 80], the dominant mechanisms compared to this study seem to be different. This is a bit surprising because in [75], manufactured materials are exactly the same as in this work. There could be many reasons for different results, mainly due to manufacturing process and device scaling. In this study the devices are in micrometer scale opposed to nanometer scale devices that are studied in referred papers. This might change the conductive filament geometry. One way to interpret the results is that the filament geometry in this work is more like two-dimensional due to larger device size and distribution of oxygen vacancies in oxide layer. QPC model and conductance quantization work better for very narrow constriction of filament that has one-dimensional type structure [74, 79], and it is possible that decreasing device dimension in $\mathrm{Ta/Ta_2O_5/TiN}$ could reveal those properties as some observed results give a hint of their existence. Especially conductance quantization in ReRAMs is promising for multi-level storage applications [82].

Pseudo-analog operation should be easier to obtain if the $i - v$ characteristics follow some clear mechanisms. In this case the dominant mechanism is confirmed to be SCLC mechanism. Because the reset behaviour varies greatly between subsequent cycles, pseudo-analog behaviour may be hard to achieve.

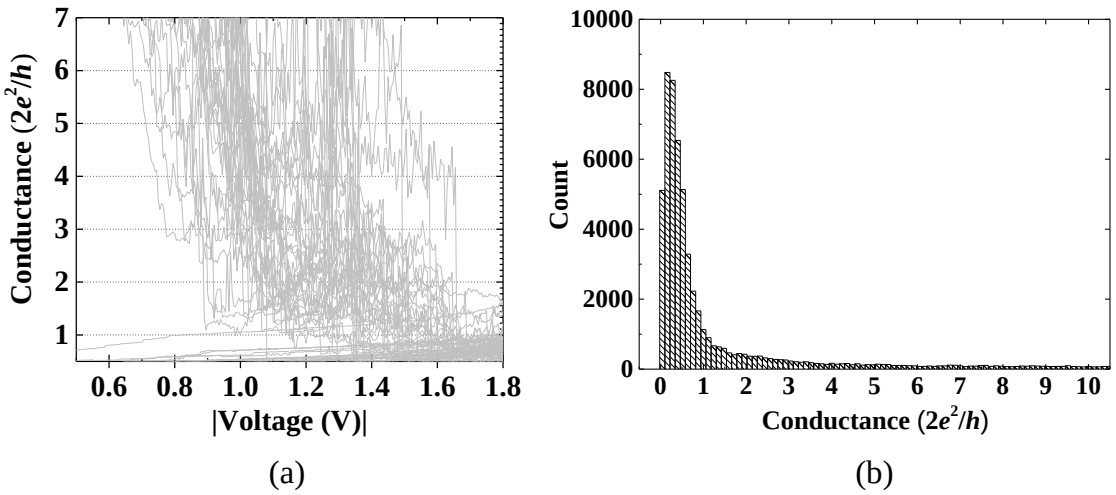

Figure 6.15: (a) Normalized conductance values of multiple reset sweeps. No clear discrete levels are visible to eye. (b) Histogram graph shows that conductance quantization phenomena is not present in the devices of this study. Measured cell is the same as in Figure 6.14.

6.3 Analog Behaviour

The main method of trying to achieve the analog operation is increasing negative stop voltage gradually and calculating the corresponding resistance values by measuring current at $-10\,\mathrm{mV}$. Pulse operation mode is not used in this study for the reasons that are presented in Section 6.1.7.

The point where reset process begins, is very well controlled and happens around $-0.5\,\mathrm{V}$ as already presented in Figure 6.9. Analog behaviour is analysed by setting ReRAM cells first into LRS and beginning the reset modulation from $-0.5\,\mathrm{V}$ up to $-1.4\,\mathrm{V}$ using $-0.1\,\mathrm{V}$ intervals. Finally, the devices are set back to HRS by increasing the reset voltage near $-2.0\,\mathrm{V}$ to induce "deep" reset. This procedure should result twelve different resistance values. If gradual increase in resistance is observed some kind of pseudo-analog operation may be possible to achieve.

Pseudo-analog characteristics of three measured devices are shown in Figure 6.16 as well as corresponding resistance values of each different stop voltages. From this figure it is very clear that there are twelve easily distinguishable intermediate resistance values between R_{OFF} and R_{ON}. Comparing the $i-v$ characteristics from this figure to the ones which are obtained from one single DC sweep (e.g. Figure 6.5), profile of the curve is very similar. There exists a gradual decreasing current slope before the region of current fluctuations around $-1.0\,\mathrm{V}$. Finally, full reset sets devices back to HRS. This suggest that the geometry of filament (that defines the $i-v$ characteristics) is decided in the

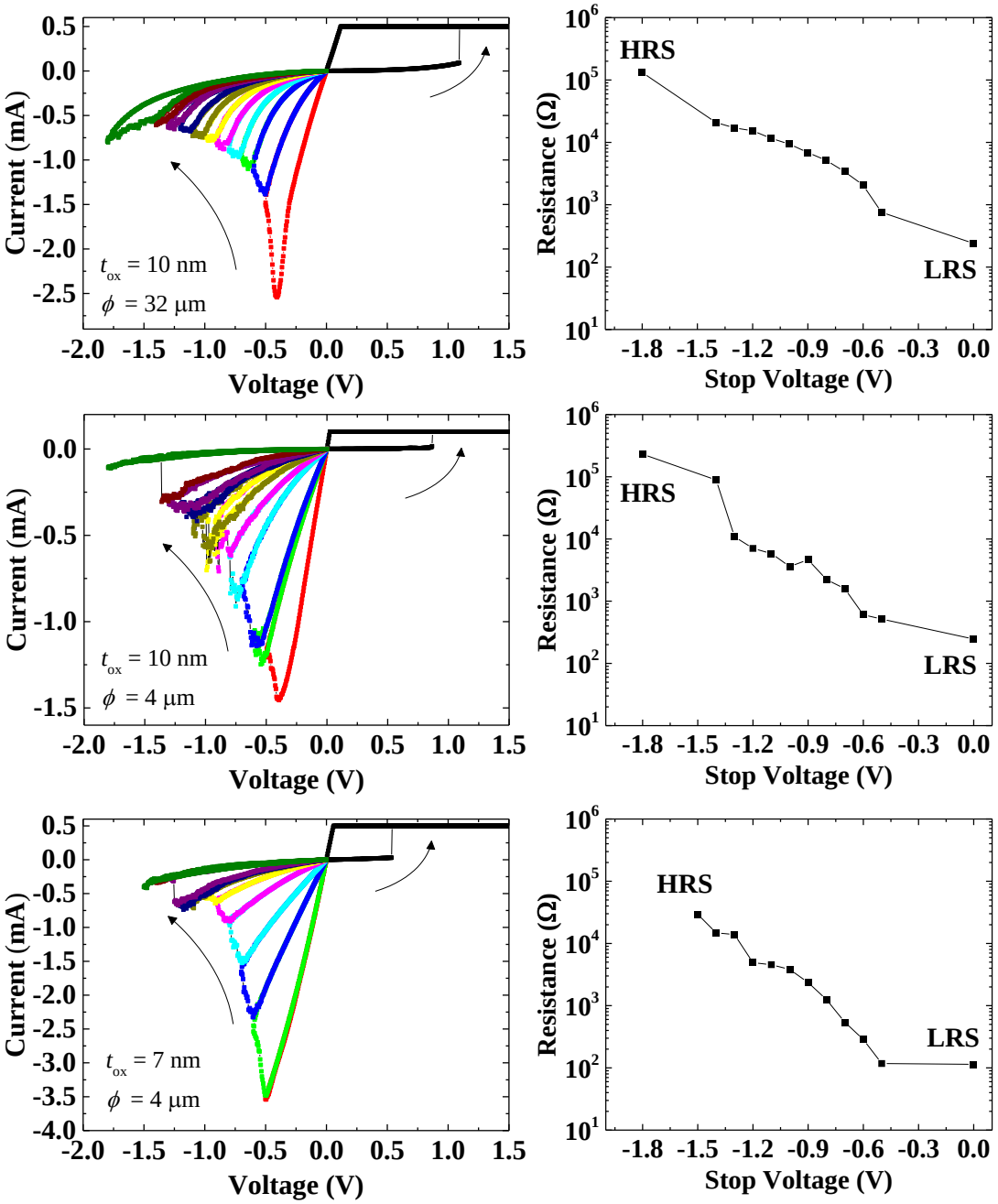

Figure 6.16: On the left, $i-v$ characteristics of three different devices are shown. Corresponding resistance values as a function of stop voltage are shown on the right side.

forming process during positive voltage DC sweep. Multiple reset sweeps do not change predetermined $i-v$ characteristics drastically.

The graphs in Figure 6.16 show the best obtained results where resistance as a function of stop voltage is *almost* decreasing in every device. However, this not the case always as there exists many cases where resistance decreases even though negative stop voltage increases. If a strictly increasing resistance values are desired, a coarser resolution of stop

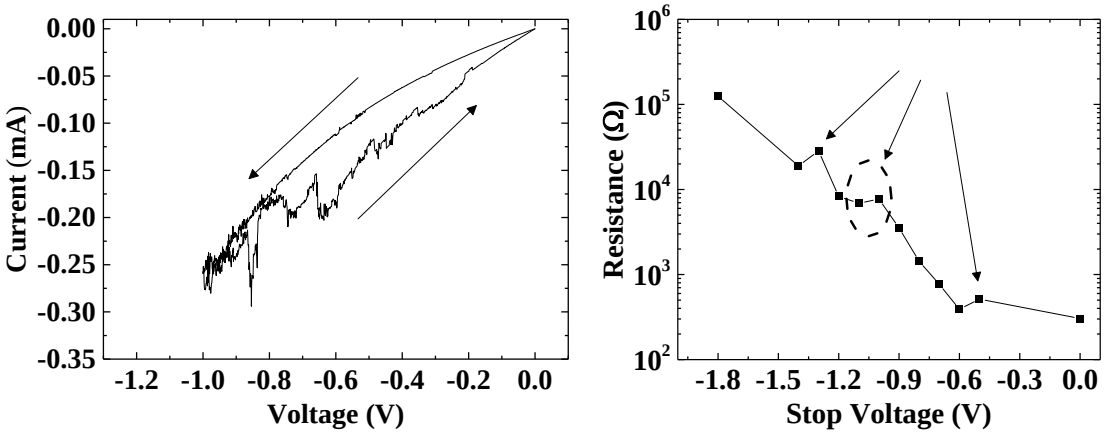

Figure 6.17: An example of one reset sweep where resistance decreases during sweep. Direction of sweep is marked by arrows. There exists three points in resistance vs. stop voltage curve where resistance decreases between subsequent cycles. The marked circle shows the point that represents the graph on the left.

voltage sweeps must be used. In this study, the interval was set to $-0.1\,\mathrm{V}$ because more frequent intervals do not provide better results. In this sense, the analog operation seems to be limited by this factor.

Analysing one particular reset sweep in Figure 6.17, demonstrates the problem if stop voltages between subsequent DC sweeps are too close too each other. Current fluctuations become too large under high electric field so that the filament might have higher conductance after DC sweep. The details of physical mechanisms behind current fluctuations are currently unknown. There seem to exists many different discrete current levels that suggest some trap to trap related tunnelling mechanisms at lower conductances. Some given explanations that are discussed in Section 6.2 could not answer this question.

Modifying resistance values in analog manner works only one way. Decreasing negative stop voltage (e.g. from $-1.0\,\mathrm{V}$ back to $-0.7\,\mathrm{V}$) does not affect resistance values and the only way to "return" back to lower resistance states is to perform a positive voltage sweep which return device back to LRS. There is no way to gradually decrease resistance values because the forming step always results a resistance value which is equivalent to R_{ON}.

A statistical analysis of resistance values between sweep-to-sweep variations is performed. It is natural to presume that analog performance gets worse because the variations of R_{OFF} and R_{ON} values is already quite significant in many measured devices.

Resistance values are collected from 30 DC sweep cycles, where one DC sweep cycle consist of one positive sweep followed by 12 reset sweeps using different stop voltages. Multiple devices were measured and some chosen results are presented in Figure 6.18. It

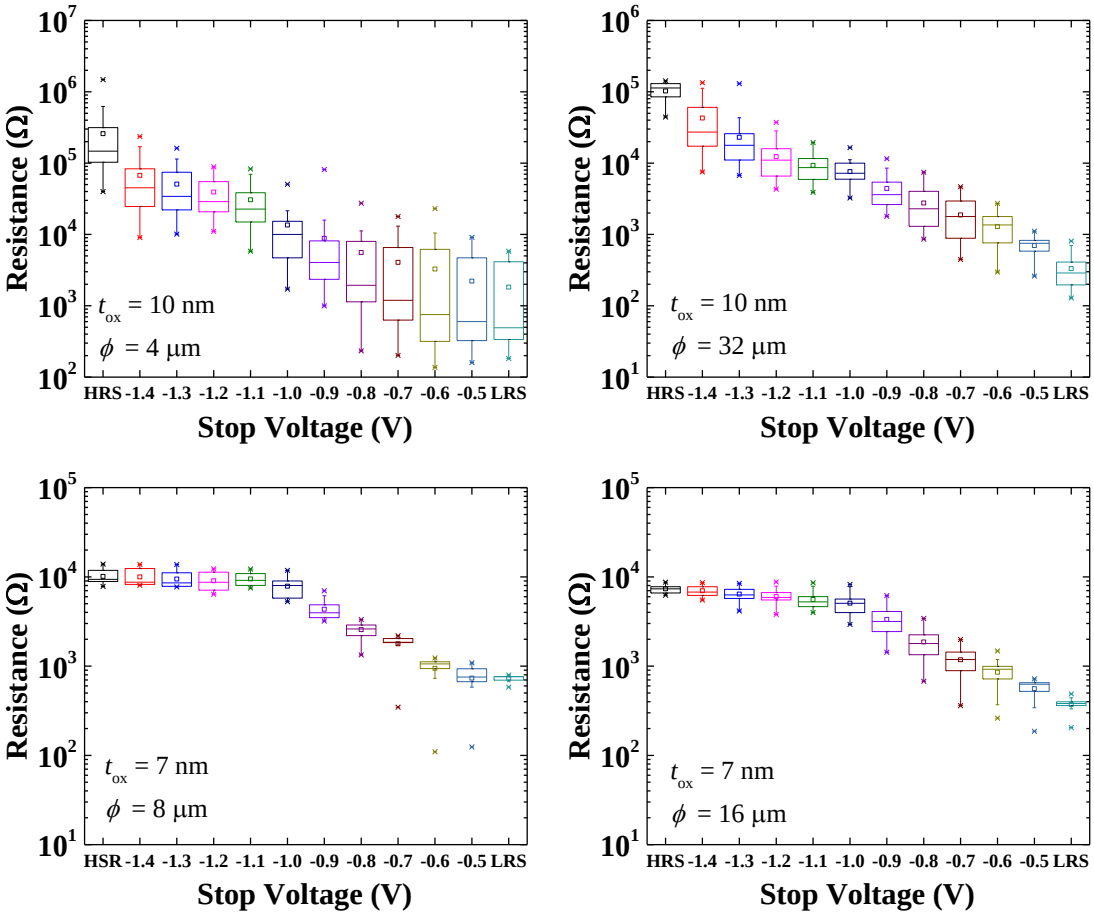

Figure 6.18: Variance of resistance values in four different devices.

is quite obvious that t_{ox} has a significant effect on distribution of intermediate resistance values. Absolute range between R_{OFF} and R_{ON} is much larger when $t_{ox} = 10\,nm$ but variance becomes larger as well. In $t_{ox} = 7\,nm$ devices the resistance range is smaller but so is the variance in intermediate resistance values. The trade-off is apparent. It also seems that "deep" reset in $t_{ox} = 7\,nm$ devices happens around $-1.1\,V$ as resistance values do not increase at higher reset voltages thus device is already in HRS.

Although average and median resistance values in Figure 6.18 seem to be increasing as a function of stop voltage, the overlap in intermediate resistance values is noticeable. This reduces the amount of reliably distinguishable resistance values in subsequent DC sweeps. The variability in analog ReRAMs' operation is a well recognized problem and attempts to reduce it have been made. For example, a double-layer ReRAM structure [73], and modified pulse operation [72], have been successfully used. Unfortunately those solutions are not possible to utilize in this study.

Individual reset sweep with multiple different stop voltages is able to produce around

10 intermediate resistance states. Statistical analysis of cells that is shown in Figure 6.18 reveals the true performance. In Figure 6.19 (a) one of the devices is chosen for further examination. Between R_{OFF} and R_{ON} there exists only three stop voltage values whose distributed resistance values do not overlap R_{OFF} or R_{ON}. On the other hand, these intermediate resistance values overlap each other so in practice only one of them could be used reliably. Figure 6.19 (b) shows all measured resistance values in individual sweeps where overlapping resistance values are apparent. Choosing one particular stop voltage e.g. $-1.0\,\mathrm{V}$, based on statistical variation shows that there exists only three repeatable resistance states in this particular ReRAM cell $\{R_{\mathrm{OFF}} \geq 40\,\mathrm{k\Omega},\ 40\,\mathrm{k\Omega} > R_{\mathrm{int}} > 1\,\mathrm{k\Omega},\ R_{\mathrm{ON}} \leq 1\,\mathrm{k\Omega}\}$. The maximum resistance in LRS and minimum resistance in HRS define a window where intermediate resistances must fit. This window as well as one particularly chosen intermediate resistance value is shown in Figure 6.19 (c). In this example there exist three suitable stop voltages $\{-0.9\,\mathrm{V},\ -1.0\,\mathrm{V},\ -1.1\,\mathrm{V}\}$ but only one of these can be chosen in single sweep.

Almost all the measured devices are able to produce two very well distinguishable resistances R_{OFF} and R_{ON}. Many devices are also able to have one intermediate resistance value between them if variance is taken into account. Multi-level storage operation is possible. It is a bit problematic to call three discrete resistance states as pseudo-analog, but on the other hand, during the single sweep there are eight to ten different resistance states present. Further discussion and suitability for the purposes of ANN is given in Chapter 8.

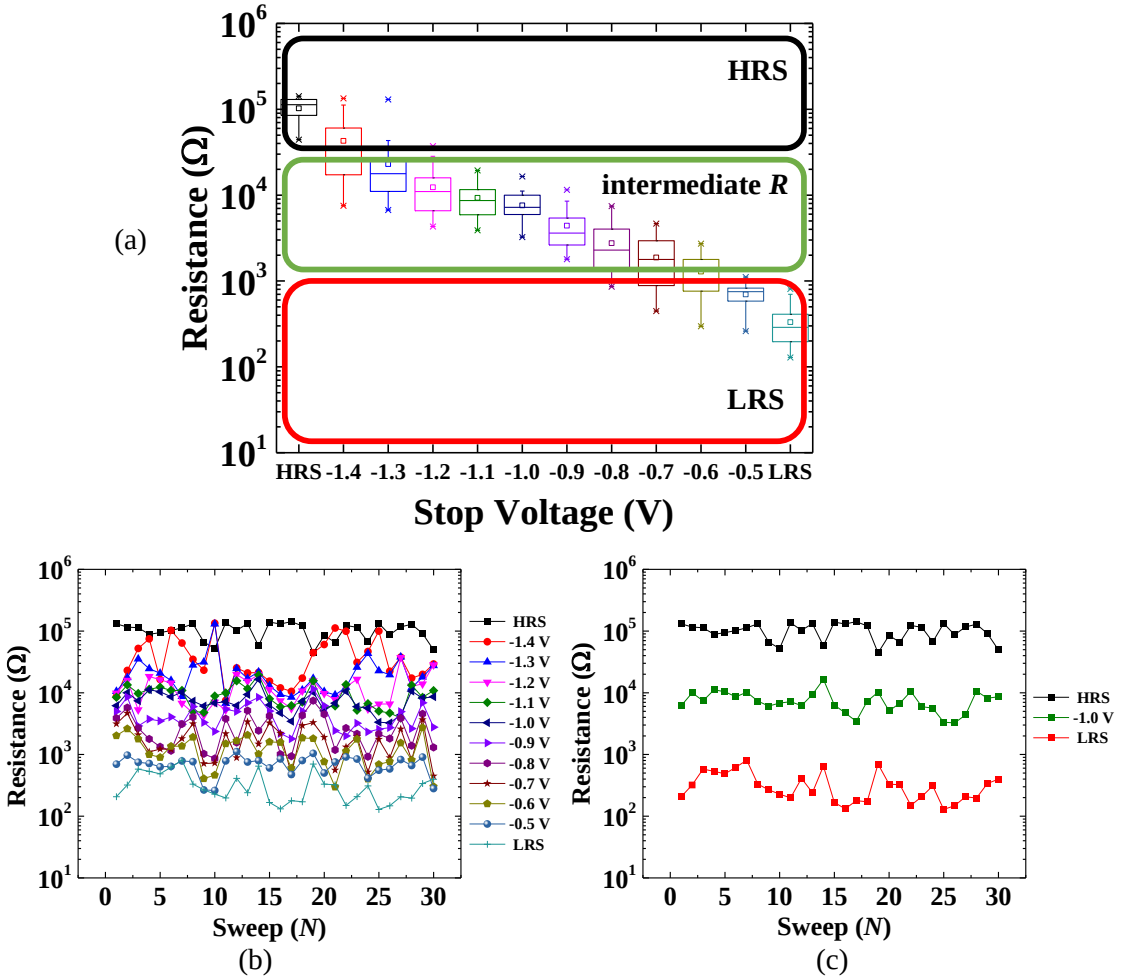

Figure 6.19: (a) Distribution of intermediate resistance values allow only one value between R_{OFF} and R_{ON}. (b) Multiple resistance values as a function of single sweep. Overlapping between different sweeps is apparent. (c) Same data as in (b) but only one chosen intermediate resistance value which is obtained using $-1.0\,\mathrm{V}$ stop voltage is shown.

7 Current Limitation Using MOSFET

Measurements of the Ta/Ta_2O_5/TiN ReRAM cell shows that the pulse operation mode is impossible without a proper current compliance circuit. This chapter studies the current compliance using MOSFET as the limiting device. One of the research methods to study analog behaviour is to investigate the resistance control by the voltage pulses considering the pulse-coupled neural network. Controlling pulse height and width should result different resistance levels in ReRAM.

7.1 Fabrication of Ta_2O_5-Based Device with MOSFET

The manufacturing process of 1T1R ReRAM is similar to the one that is presented in Chapter 5. Fabrication steps that were different are explained here.

The initial samples were provided by Kyushu Institute of Technology. Samples were cut from silicon wafer to be 15 mm × 15 mm in size and they already had n-type MOSFETs embedded. The aluminium metal pads that were connected to MOSFET's source and drain acted as the bottom electrode of ReRAM and thus no deposition of bottom electrode material was needed.

In the first photolithography step, HDMS was applied on top of samples for better adhesion of actual photoresist. Photoresist application was OFPR-800-54CP that has a higher coefficient of viscosity. Different spin coater settings that are shown in Table 7.1 were used because of that. Samples were baked at 110 °C for 90 s after coating.

Table 7.1: Spin coater settings.

Time (s)	Rotation speed (RPM)
0 - 10	0 - 500
10 - 20	500 - 2000
20 - 30	2000 - 0

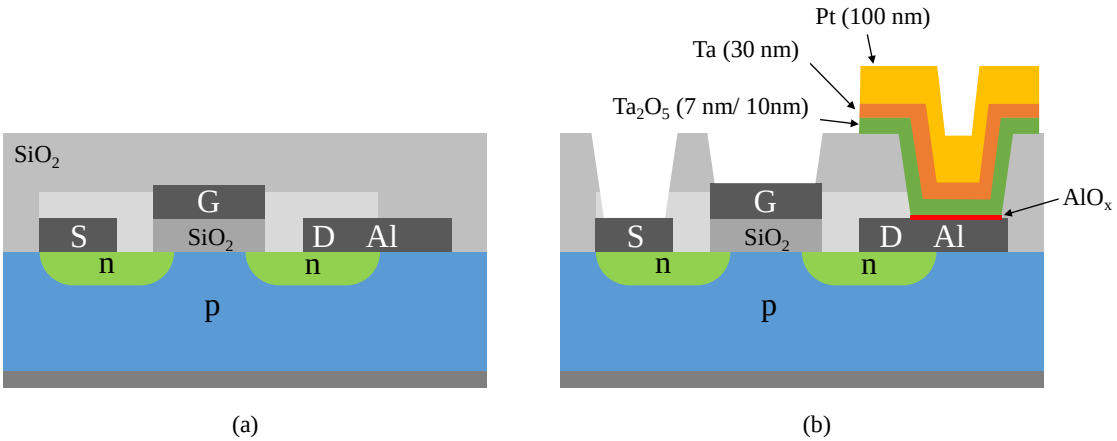

Figure 7.1: (a) Initial sample with pre-built n-type MOSFET. (b) ReRAM device manufactured on top of MOSFET. AlO_x layer is chemically formed under Ta_2O_5 layer.

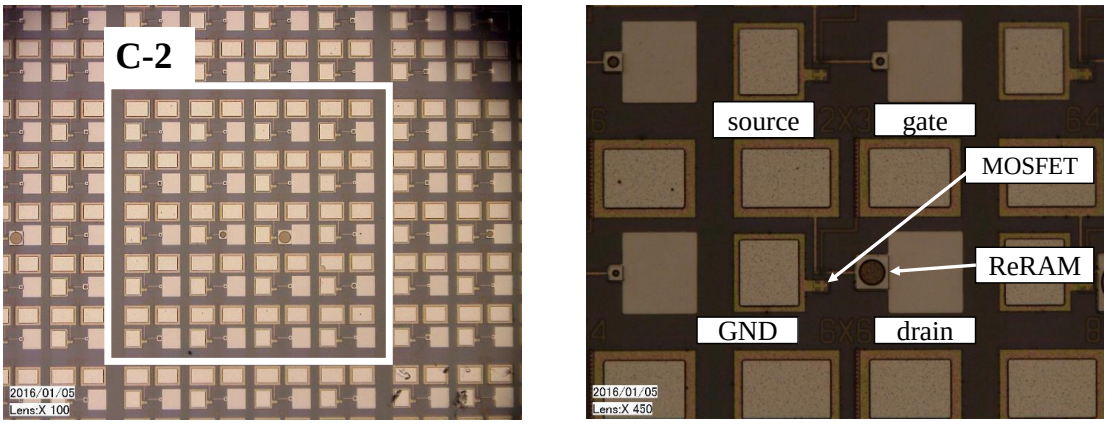

Figure 7.2: (a) Chip is divided into 36 identical sections, $\{1,2,3,4,5,6\} \times \{A,B,C,D,E,F\}$. A C-2 section is shown. Each section has 16 ReRAM devices with different ϕ. (b) Magnified image of $\phi = 32\,\mu m$ ReRAM cell. Locations of measurement pads, MOSFET and ReRAM cell are marked.

Exposure needed higher UV-light dose because of the different photoresist application OFPR-800-54CP. The exposure time was calculated to be equivalent with dose of 200 mJ. NMD-3 was used to develop samples after photolithography following by post-bake for 5 min at 135 °C.

The photolithography step exposed opening patterns that were etched to reveal the bottom electrode material, MOSFET's Al pads. RIE was used, but this time with longer etching time, 25 min 39 s because the samples had initially manufactured thicker SiO_2 layer. The bottom electrode patterns that define the ReRAM area had following hole sizes

$$\phi = \{4\,\mu m, 6\,\mu m, 8\,\mu m, 12\,\mu m, 16\,\mu m, 32\,\mu m, 64\,\mu m\}.$$

The rest of the manufacturing process was exactly the same as explained in Chapter 5. This time only two chips with the following oxide thicknesses were manufactured, $t_{ox} = \{7\,\text{nm},\ 10\,\text{nm}\}$. Those values were chosen because they were proven to work in $Ta/Ta_2O_5/TiN$ ReRAM cell. The cross section of $Ta/Ta_2O_5/AlO_x/Al$ ReRAM cell is shown in Figure 7.1 and microscope images in Figure 7.2.

7.2 Measurement and Characteristics

The biggest difference compared to $Ta/Ta_2O_5/TiN$ ReRAM cell is the absence of TiN in bottom electrode. Etching opened holes directly to MOSFET aluminium pads which act as the bottom electrodes in 1T1R ReRAM. There exists a very thin chemically formed AlO_x layer under Ta_2O_5 which may be the biggest reason why the $i-v$ characteristics change significantly compared to measurements in Chapter 6 even though the oxide and top electrode materials are same. Another big difference comes from MOSFET current limitation whose schematic and voltage nodes are shown in Figure 7.3 (a) and $i-v$ characteristics in Figure 7.3 (b).

The measurement probes were set as follows: one waveform generator/fast measurement unit (WFGMU) controlled the drain voltage v_D, and another WFGMU controlled MOSFET's gate voltage, v_G. MOSFET's bulk contact was connected to common ground together with source voltage, v_S. Figure 7.3 (b) shows the difference of the $i-v$ characteristics compared to $Ta/Ta_2O_5/TiN$ ReRAM cell. The MOSFET's v_G decides the current compliance limit. If a strict control of some current compliance e.g. $100\,\mu\text{A}$ is required, MOSFET's $i-v$ characteristics need to be investigated. MOSFET characterization is not the focus in this study so measurements were not made. However, results from a previous work [83], that used exactly the same MOSFET are used as a reference. Some of the MOSFET's $i-v$ characteristics are presented in Figure 7.4.

MOSFET's $i-v$ curves show that $100\,\mu\text{A}$ current compliance, which is commonly used in this study, is achieved by setting v_G to $1.85\,\text{V}$. Proper current compliance is necessary for pulse operation when pulse amplitude has an effect on desired resistance level. Sometimes strict control of current compliance is used for modulating R_{ON} values in LRS. This method is successfully used in [84], for instance. One of the important requirements for 1T1R strucutre is the low channel resistance of MOSFET. If the channel resistance is relatively high compared to ReRAM cell's R_{ON}, it will be dominating thus effectively

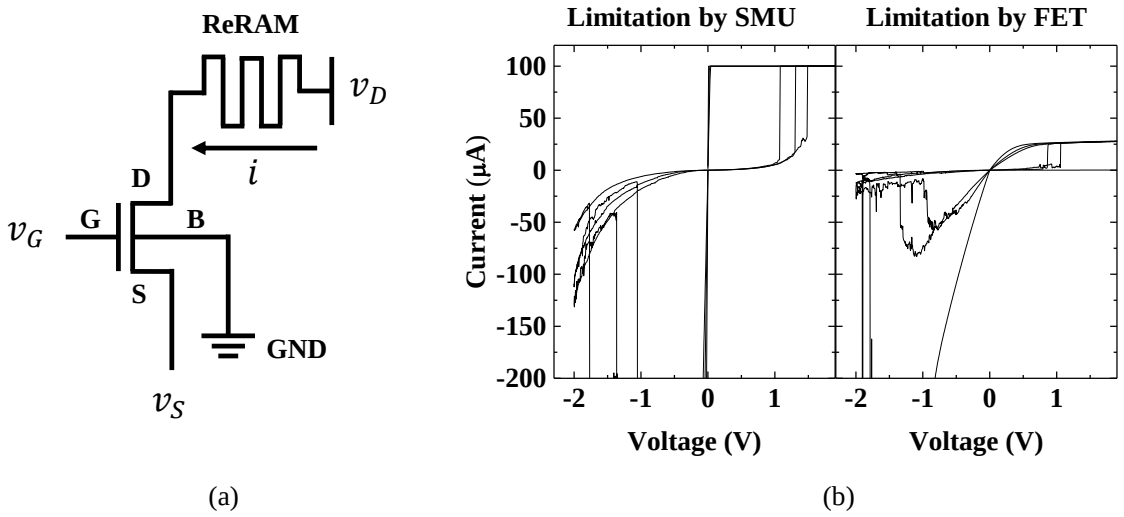

Figure 7.3: (a) Schematic of 1T1R structure. (b) Current compliance by MOSFET reveals the different $i - v$ characteristics compared to limitation by B1500A's SMU.

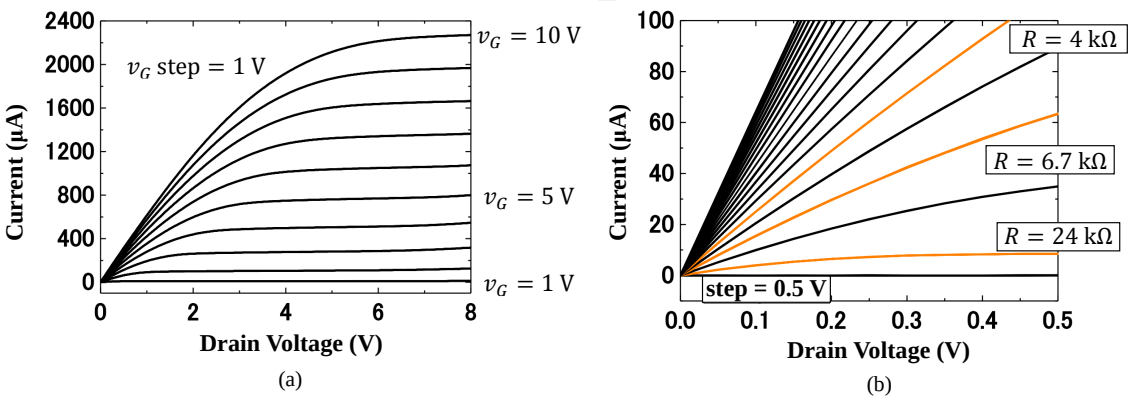

Figure 7.4: (a) MOSFET current as a function of drain voltage. Gate voltage, v_G is increasing in 1 V steps. (b) Three calculated MOSFET channel resistances (shown in orange) in triode region where corresponding $v_G = \{1\,\text{V}, 2\,\text{V}, 3\,\text{V}\}$. Slightly adapted from [83].

reducing $R_{\text{OFF}}/R_{\text{ON}}$. Low channel resistance is achieved by increasing v_G. Typical read voltage of ReRAM could be 10 mV which fill fall into triode region of MOSFET. In this case the MOSFET's channel resistance is 1.6 kΩ when $v_G = 10\,\text{V}$. Channel resistance is definitely higher than $Ta/Ta_2O_5/TiN$ cell's R_{ON}.

One of the first observations in 1T1R cell, $Ta/Ta_2O_5/AlO_x/Al$ is the different $i - v$ characteristics in contrast to $Ta/Ta_2O_5/TiN$ cell. The initial forming voltages and resistances are higher as well as other operating points are different. The observations that are made in Chapter 6, may not apply to analog operation in $Ta/Ta_2O_5/AlO_x/Al$ device. Two examples of $i - v$ characteristics are given in Figure 7.5. Even in $t_{\text{ox}} = 7\,\text{nm}$ devices, the initial forming voltage is about three times higher compared to $Ta/Ta_2O_5/TiN$

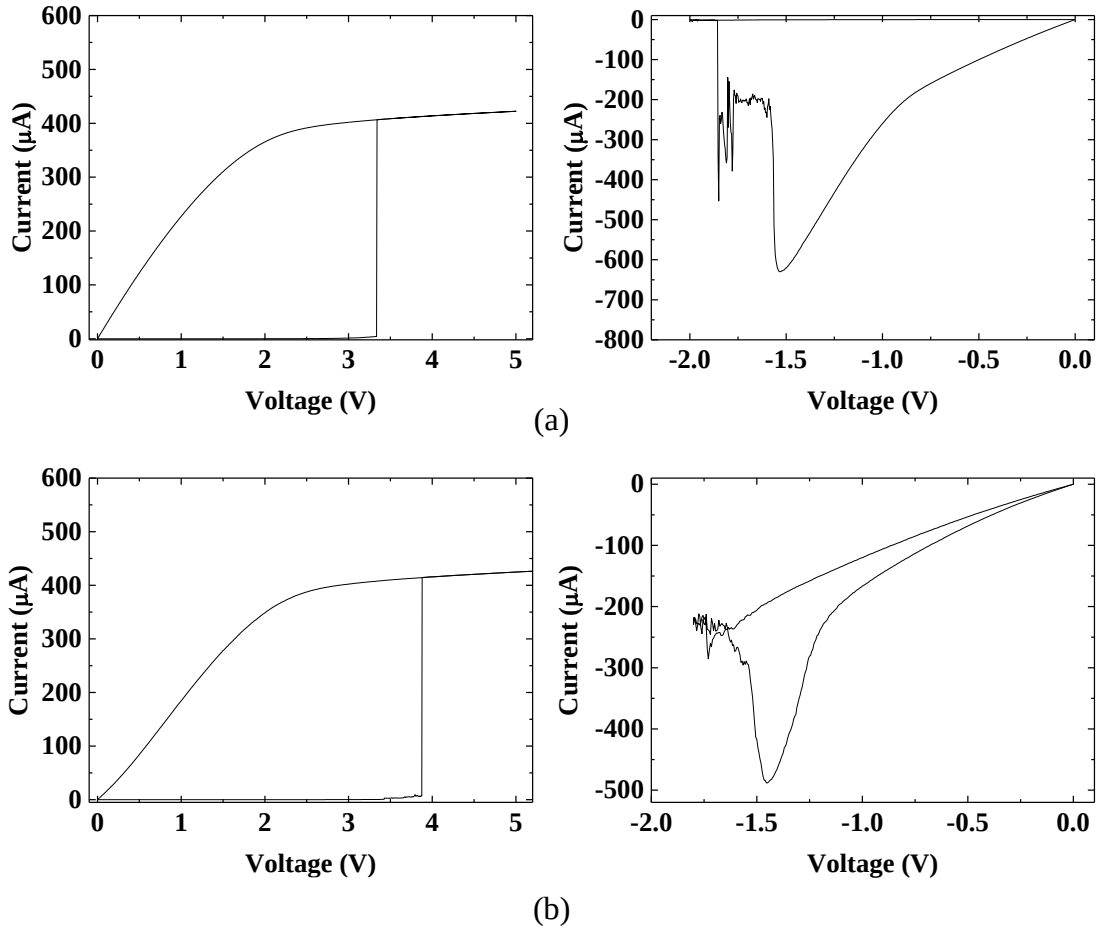

Figure 7.5: Initial forming and following reset sweeps of $Ta/Ta_2O_5/AlO_x/Al$ cell. Device dimensions are (a) $t_{ox} = 7\,nm$, $\phi = 12\,\mu m$ and (b) $t_{ox} = 10\,nm$, $\phi = 6\,\mu m$.

cell. Initial resistance is too high to measure with B1500A. Also the reset behaviour is different. Reset occurs at around $-1.5\,V$ and the gradual decreasing slope is missing. This kind of behaviour is not desirable if analog operation by reset voltage modulation is considered.

Figure 7.5 (b) shows the reset process in $t_{ox} = 10\,nm$ device. There is no "deep" reset present. If the reset sweep is continued further, all $t_{ox} = 10\,nm$ devices brake down before switching to HRS happens. This behaviour is very similar to $Ta/Ta_2O_5/TiN$ $t_{ox} = 20\,nm$ cells. The device is not able to support switching as further switching behaviour is not observed in subsequent positive or negative voltages sweeps.

Unfortunately, resistive switching in proposed $Ta/Ta_2O_5/AlO_x/Al$ ReRAM cells is very unreliable regardless of device dimensions. Most of the time the devices brake down after few sweeps and many of the measured devices are not able to produce any switching behaviour at all. Figure 7.6 shows obscure $i - v$ characteristics that is present in

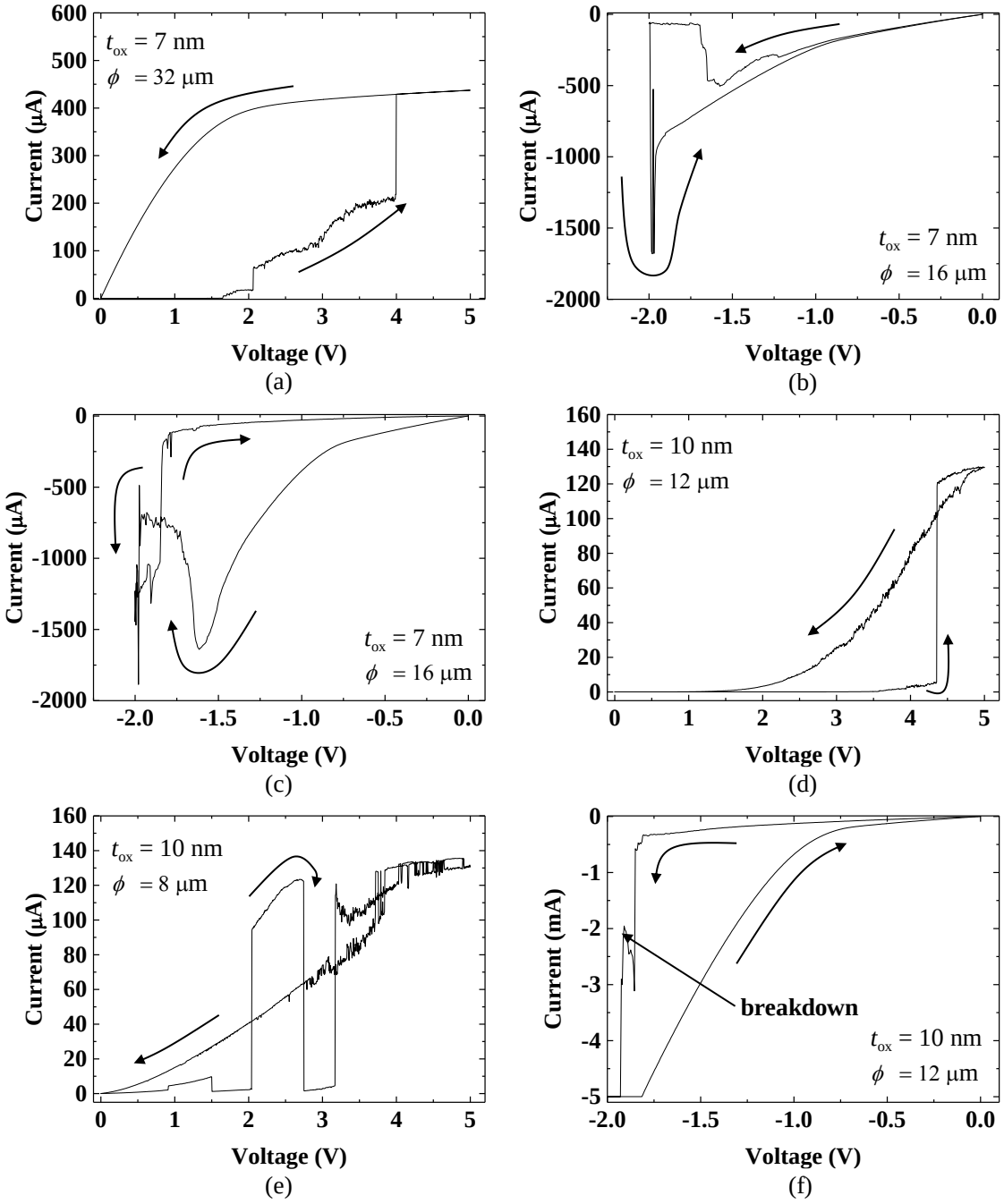

Figure 7.6: Examples of current behaviour on different voltage sweeps.

$Ta/Ta_2O_5/AlO_x/Al$ cells.

In Figure 7.6 (a) forming process happens in steps. The first jump to slightly conducting state occurs at 2 V although a complete switch to LSR occurs at 4 V. The reason for random fluctuations before abrupt set is probably related to AlO_x layer. Figure 7.6 (b) and (c) show the abnormal behaviour in reset sweep. For example, Figure 7.6 (b) has a clear reset step at −1.7 V but suddenly a jump back to lower resistance state happens at

$-2.0\,\mathrm{V}$. Similar behaviour is observed in Figure 7.6 (c) where abrupt jumps to different resistance levels happen under high electric fields. In this case the device is able to switch back to HRS but the reset process is very hard to predict.

Two examples of $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ set processes are also presented. Figure 7.6 (d) has a clear abrupt set around $4.3\,\mathrm{V}$ but device is not able to sustain LRS. In fact the device is in some undefined state between LSR and HRS as further resistive switching could not be produced. Situation is similar in Figure 7.6 (e) where multiple jumps between different conductive states are observed in single DC sweep. Again, further resistive switching could not be produced. Figure 7.6 (e) shows a very common reset behaviour in $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ devices where breakdown occurs at higher voltages before full reset happens.

The conclusion is that proposed $\mathrm{Ta/Ta_2O_5/AlO_x/Al}$ ReRAM cells are able to produce resistive switching but the current manufacturing process does not provide reliable devices. This is the reason why pulse operation is left out from study. Most of the ReRAM test structures are not able to produce any kind of resistive switching behaviour. The current behaviour resembles electrical noise in all voltage levels. The few working devices are able to produce resistive switching in subsequent sweeps only few times before breakdown occurs. The exact reason for this is unknown but breakdowns are very likely related to manufacturing process. Probably the surface of Al layer is very rough and has irregularities in nanometer range. This becomes a problem when oxide layer is very thin which is exactly the case in proposed structure. However, the resistive switching is already unreproducible at $t_{\mathrm{ox}} = 10\,\mathrm{nm}$ so increasing oxide thickness is unlikely a solution. There are ways to achieve almost perfect and flat surfaces in manufacturing process such as *chemical mechanical planarization* (CMP). If TiN layer is not built on top of MOSFET via sputtering, CMP or some other process may be the only way to solve switching problems.

8 Conclusions

Resistive random access memory could be the next breakthrough technology that competes with traditional non-volatile memories such as Flash. There are still many obstacles to be solved before reliable commercial products will be widely adopted. There is in-depth research carried out by private companies as well as scientific communities which shows that the potential of ReRAM devices is very well recognized.

ReRAMs can be used to store multiple bits which is comparable to Flash's multi-level cell (MLC) operation. Because the switching operation of ReRAM relies on the resistance change in oxide material the multiple resistance states between on-resistance, R_{ON} and off-resistance, R_{OFF} should be investigated. This kind of operation is not only limited to some certain discrete resistance levels but analog behaviour could also be a possible option. In theory, a ReRAM device can be said to be equivalent to a memristor which is a passive analog component. In practice, the operation is fully analog if continuous range of resistance values between R_{OFF} and R_{ON} can be utilized. Typically the device has multiple discrete resistance levels that are possible to achieve practically and thus it is more appropriate to call the operation pseudo-analog.

The term ReRAM may be a bit misleading because it suggests that such components can be only used in memory technologies. ReRAMs have been proposed to do computations similarly than traditional CMOS logic gates. There are also theoretical proposals and proof-of-concepts about reconfigurable architectures of analog integrated circuits. However, the most common approach is to manufacture ReRAM cells in crossbar like memory array structure. An interesting idea is to utilize the crossbar structure to imitate neural network. Artificial neural networks (ANN) have many implementations in software and hardware level but as it turns out, one way to construct ANN is a crossbar array. In this case, the ReRAM device acts as a synaptic strength w in a matrix multiplication formalism $b_j = \sigma(\sum a_i w_{i,j})$. Synaptic strength is the resistance of ReRAM cell in the equation. More resistance states allows more complex computations in network which is the reason

why pseudo-analog operation is studied in this thesis keeping ANN in mind.

The proposed ReRAM structure is Ta/Ta_2O_5/TiN. The measurements were conducted using semiconductor parameter analyser which provided the necessary $i - v$ sweep and pulse operation modes for ReRAM cell's characterization. The varying manufactured parameters of ReRAM cells are oxide thickness and cell area. Three different test chips were manufactured where each of them had identical test structures. The test chips had the following oxide thicknesses, t_{ox}: 7 nm, 10 nm and 20 nm and within each chip, cells with following diameters, ϕ were measured: 4 µm, 8 µm, 16 µm, 32 µm and 64 µm. The purpose of the different test structures was to see how different dimensions affect on ReRAM's $i-v$ characteristics and pseudo-analog operation.

One of the first observations is that the ReRAM cells with $t_{\text{ox}} = 20\,\text{nm}$ could not sustain resistive switching before breakdown. There is hysteresis memory characteristics present during initial positive DC sweep but during reset sweep the device is not able to switch ReRAM back to HRS from LRS. This is the main reason why analog operation as well as many other important performance metrics are missing for that particular oxide thickness. Two other test chips where $t_{\text{ox}} = 7\,\text{nm}$ and $t_{\text{ox}} = 10\,\text{nm}$ are able to achieve hysteresis characteristics in $i - v$ plane. However, the initial resistance becomes too low for $t_{\text{ox}} = 7\,\text{nm}$ ReRAM cells when device area is increased. Large cells where $\phi = 64\,\mu\text{m}$ do not show resistive switching behaviour.

The detailed measurement results are given in Chapter 6 and those are not repeated here but the pseudo-analog behaviour is discussed. One of the important factors is the amount of intermediate resistance states, R_{int} between R_{OFF} and R_{ON}. Usually, R_{int} occurs in discrete resistance levels, and therefore $R_{\text{OFF}}/R_{\text{ON}}$ is an important factor. It turns out that t_{ox} is the dominant factor that defines the $R_{\text{OFF}}/R_{\text{ON}}$. A typical ratio for $t_{\text{ox}} = 7\,\text{nm}$ cells was around 10 while for $t_{\text{ox}} = 10\,\text{nm}$ it was around 50. The observation here is that increasing oxide thickness will increase $R_{\text{OFF}}/R_{\text{ON}}$ exponentially which should be promising for pseudo-analog operation. However, this comes with a trade-off because the observed variance of R_{OFF} and R_{ON} values between subsequent sweeps became larger when oxide layer is thicker. Large variance in resistance values may be problematic if strict control of discrete resistance levels is desired.

Analysing the $i - v$ characteristics showed some common behaviour that is observed in all devices. Typically, all devices had some sort of leakage current present in HRS that increased exponentially as a function of voltage. The exponential leakage is proportional

to t_{ox} and ϕ. Sometimes exponential current increase in LRS is observed just before reset process start to occur. The behaviour is interpreted as SCLC mechanism. In some cases there are clear conductance levels present during the progressive reset phase. Other similar studies found in literature explain such behaviour using conductance quantization phenomenon and QPC-model which is linked with conductance quantization. An attempt to see if $Ta/Ta_2O_5/TiN$ structure would follow those models is made. The fitting of the model is not very well consistent with measured results which is somehow surprising as other studies have used the same materials, although the device dimension and manufacturing processed were different. Especially conductance quantization is interesting because it should provide repeatable resistance values in integer multiplies of G_0. This could be used in practice for multi-level memory operation.

All the devices that are able to sustain resistive switching are also analysed for analog behaviour. The main method is to modulate the reset voltage by increasing the negative stop voltage value because there is a gradual decrease in the current observed during the reset sweep. The results obtained from a single reset sweep with modulation of stop voltages by using $-0.1\,\mathrm{V}$ intervals are very promising. Most of the times the resistance values are increasing although between subsequent sweeps there are cases where the interval is too small to induce any change in resistance or it may even decrease the resistance value. If the resistance has to be strictly increasing as a function of increasing negative stop voltage, the only way to achieve this is using coarser intervals of stop voltages. It turned out that between R_{OFF} and R_{ON} there are usually eight to ten resistance values. These results can be considered good as there do not exist many other studies that have achieved considerably more intermediate resistance values. The true performance is revealed if those values can be reproduced easily.

To fully understand the pseudo-analog characteristics, the statistical variation between multiple reset sweeps is also analysed. The 30 subsequent cycles where a switch from HRS to LRS is performed between modulation of stop voltages. The device size does not have any significant effect on variability of intermediate resistance states. The oxide thickness is the main factor as thinner oxide layers seem to reduce the variance which is observed by comparing the results that are obtained from $t_{ox} = 7\,\mathrm{nm}$ and $t_{ox} = 10\,\mathrm{nm}$ cells. If only memory operation is considered, the ReRAM cell is able to store three resistance values that can be presented with two bits. The intermediate resistance level can be achieved by multiple different stop voltages but only one could be used during single reset sweep.

Otherwise the variance between subsequent sweeps overlap each other thus reading the resistance could be misinterpret.

Strict current control is necessary to prevent the breakdown of ReRAM cell. It turned out that B1500A was not enough to provide sufficient compliance current because the switching of ReRAM cell is too fast. This led to many broken devices during measurements and made the pulse operation measurements impossible to conduct. It was mentioned in Chapter 1 that investigation of resistance control by voltage pulses keeping pulse-coupled neural network in mind should be considered. Another set of test chips were manufactured that had an embedded MOSFET in series with ReRAM. A MOSFET can provide strict control of compliance current by modulating the gate voltage. It is also fast enough to allow pulse operation with high frequency of voltage pulses. Unfortunately the current manufacturing process is not suitable to produce devices that are able to sustain resistive switching for multiple cycles. Most of the devices were broken initially and the rest were broken after few sweeps. No useful data is obtained to analyse $i-v$ characteristics or pulse operation for pseudo-analog purposes. The reason for this is probably due to roughness of MOSFET's Al pads that acted as bottom electrode. There is also chemically formed AlO_x layer that is present between bottom electrode and oxide layer which changes the ReRAM cell's operation significantly.

In practical applications, the pulse programming is the only way to achieve fast operation and low power consumption. It can be said that currently the proposed ReRAM cell is not suitable for any practical application without proper current control. However, a more difficult question is whether or not pseudo-analog operation is present in devices. The statistical analysis revealed that, if overlapping between resistance values is not allowed there exists only three distinguishable resistance levels. Three discrete resistance levels can be hardly called pseudo-analog operation. On the other hand, the single reset sweep revealed eight to ten different distinguishable resistance levels which could be interpret as pseudo-analog.

The motivation of this study was to see if Ta_2O_5-based ReRAM cell is suitable for pseudo-analog operation considering ANN. Currently the proposed cell is suitable for binary weighted synapses instead of pseudo-analog synapses regardless of device dimensions. In that sense, some refinements should be made to utilize the potential of multiple resistance states. Controlling the cell dimensions, such as oxide thickness's uniformity, is crucial for predictable operation as well as minimizing variance. This would require

other manufacturing methods like ALD instead of RF sputtering. Strict control of manufacturing process is important because otherwise the variability between adjacent cells is too large. Choice of materials seems to be good for pseudo-analog operation because single device is able to produce multiple resistance states although reproducibility is still a problem. Another concern is the current limitation. $Ta/Ta_2O_5/TiN$ structure can not limit the current without external components, which is the reason why 1T1R structure is the choice for further studies. It also enables pulse operation mode that is crucial when pulse-coupled ANN is considered.

The requirements for pseudo-analog operation may depend on the application where proposed ReRAM cell is used. For example, some stochastic computing could use the statistical variation instead of strict resistance values. This thesis does not study the effect on variability of synapse strengths in ANN and it could be possible that some specific applications such as optimization problems may not suffer from variance. There actually exist a subtype of ANN called *stochastic neural network* [85], and it is possible to give stochastic weights to network's synapses. One example of a stochastic neural network is *Boltzmann machine*. Boltzmann machine is a recurrent neural network instead of the feed-forward type network that is presented in Section 4.4. Stochastic computations may be much more complicated to perform with crossbar arrays. This kind of analysis could be the motivation for further studies where this thesis serves as a basis of utilizing Ta_2O_5-based ReRAM cells in such cases.

Bibliography

[1] L. Chua, "Memristor–The missing circuit element," *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971.

[2] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found.," *Nature*, vol. 453, no. 7191, pp. 80–3, 2008.

[3] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, "Metal–oxide RRAM," *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.

[4] L. Chua and S. M. Kang, "Memristive devices and systems," *Proceedings of the IEEE*, vol. 64, no. 2, pp. 209–223, 1976.

[5] L. Chua, "Resistance switching memories are memristors," *Applied Physics A: Materials Science and Processing*, vol. 102, no. 4, pp. 765–783, 2011.

[6] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, p. 500, 1952.

[7] S. Vongehr and X. Meng, "The missing memristor has not been found," *Scientific Reports*, vol. 5, p. 11657, 2015.

[8] J. F. Gibbons and W. E. Beadle, "Switching properties of thin NiO films," *SolidState Electronics*, vol. 7, no. 11, pp. 785–790, 1964.

[9] G. Dearnale, a. M. Stoneham, and D. V. Morgan, "Electrical phenomena in amorphous oxide films," *Rep. Prog. Phys.*, vol. 33, pp. 1129–1191, 1970.

[10] S. Seo, M. J. Lee, D. H. Seo, E. J. Jeoung, D.-S. Suh, Y. S. Joung, I. K. Yoo, I. R. Hwang, S. H. Kim, I. S. Byun, J.-S. Kim, J. S. Choi, and B. H. Park, "Reproducible

resistance switching in polycrystalline NiO films," *Applied Physics Letters*, vol. 85, no. 23, pp. 5655–5657, 2004.

[11] C. Rohde, B. J. Choi, D. S. Jeong, S. Choi, J.-S. Zhao, and C. S. Hwang, "Identification of a determining parameter for resistive switching of TiO2 thin films," *Applied Physics Letters*, vol. 86, no. 26, pp. –, 2005.

[12] W. Zhuang, W. Pan, B. Ulrich, J. Lee, L. Stecker, A. Burmaster, D. Evans, S. Hsu, M. Tajiri, A. Shimaoka, K. Inoue, T. Naka, N. Awaya, A. Sakiyama, Y. Wang, S. Liu, N. Wu, and A. Ignatiev, "Novel colossal magnetoresistive thin film nonvolatile resistance random access memory (RRAM)," *Digest. International Electron Devices Meeting,*, pp. 193–196, 2002.

[13] I. G. Baek, M. S. Lee, S. Seo, M. J. Lee, D. H. Seo, D. S. Suh, J. C. Park, S. O. Park, H. S. Kim, I. K. Yoo, U. I. Chung, and J. T. Moon, "Highly scalable non-volatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses," in *Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International*, pp. 587–590, 2004.

[14] A. Ascoli, F. Corinto, V. Senger, and R. Tetzlaff, "Memristor model comparison," *IEEE Circuits and Systems Magazine*, vol. 13, no. 2, pp. 89–105, 2013.

[15] M. D. Pickett, D. B. Strukov, J. L. Borghetti, J. J. Yang, G. S. Snider, D. R. Stewart, and R. S. Williams, "Switching dynamics in titanium dioxide memristive devices," *Journal of Applied Physics*, vol. 106, no. 7, pp. 1–6, 2009.

[16] C. A. Mead, "Anomalous capacitance of thin dielectric structures," *Physical Review Letters*, vol. 6, no. 10, pp. 545–546, 1961.

[17] D. B. Strukov and R. S. Williams, "Exponential ionic drift: Fast switching and low volatility of thin-film memristors," *Applied Physics A: Materials Science and Processing*, vol. 94, no. 3, pp. 515–519, 2009.

[18] J. P. Strachan, D. B. Strukov, J. Borghetti, J. J. Yang, G. Medeiros-Ribeiro, and R. S. Williams, "The switching location of a bipolar memristor: chemical, thermal and structural mapping.," *Nanotechnology*, vol. 22, no. 25, p. 254015, 2011.

[19] D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim, and C. S. Hwang, "Atomic structure of conducting

nanofilaments in TiO2 resistive switching memory.," *Nature nanotechnology*, vol. 5, no. 2, pp. 148–153, 2010.

[20] R. Waser, R. Dittmann, C. Staikov, and K. Szot, "Redox-based resistive switching memories nanoionic mechanisms, prospects, and challenges," *Advanced Materials*, vol. 21, no. 25-26, pp. 2632–2663, 2009.

[21] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng, "Recent progress in resistive random access memories: Materials, switching mechanisms, and performance," *Materials Science and Engineering R: Reports*, vol. 83, no. 1, pp. 1–59, 2014.

[22] A. Sawa, "Resistive switching in transition metal oxides," *Materials Today*, vol. 11, no. 6, pp. 28–36, 2008.

[23] M. Arita, A. Takahashi, Y. Ohno, A. Nakane, A. Tsurumaki-Fukuchi, and Y. Takahashi, "Switching operation and degradation of resistive random access memory composed of tungsten oxide and copper investigated using in-situ TEM," *Scientific Reports*, vol. 5, no. August, p. 17103, 2015.

[24] H. Kondo, M. Arita, T. Fujii, H. Kaji, M. Moniwa, T. Yamaguchi, I. Fujiwara, M. Yoshimaru, and Y. Takahashi, "The observation of "conduction spot" on NiO resistance random access memory," *Japanese Journal of Applied Physics*, vol. 50, no. 8R, p. 081101, 2011.

[25] M. Kudo, M. Arita, Y. Ohno, and Y. Takahashi, "Filament formation and erasure in molybdenum oxide during resistive switching cycles," *Applied Physics Letters*, vol. 105, no. 17, pp. 2014–2017, 2014.

[26] M. Kudo, M. Arita, Y. Takahashi, K. Ohba, M. Shimuta, and I. Fujiwara, "Visualization of conductive filament during write and erase cycles on nanometer-scale ReRAM achieved by in-situ TEM," in *2015 IEEE International Memory Workshop (IMW)*, pp. 85–88, 2015.

[27] H. Y. Lee, P. S. Chen, T. Y. Wu, Y. S. Chen, C. C. Wang, P. J. Tzeng, C. H. Lin, F. Chen, C. H. Lien, and M. J. Tsai, "Low power and high speed bipolar switching with a thin reactive Ti buffer layer in robust HfO2 based RRAM," *Technical Digest - International Electron Devices Meeting, IEDM*, pp. 3–6, 2008.

[28] Z. Fang, H. Y. Yu, X. Li, N. Singh, G. Q. Lo, and D. L. Kwong, "HfOx/TiOx/HfOx/TiOx multilayer-based forming-free RRAM devices with excellent uniformity," *IEEE Electron Device Letters*, vol. 32, no. 4, pp. 566–568, 2011.

[29] Y. Kang, M. Verma, T. Liu, and M. K. Orlowski, "Formation and rupture of multiple conductive filaments in a Cu/TaOx/Pt device," *ECS Solid State Letters*, vol. 1, no. 5, pp. Q48–Q50, 2012.

[30] S. Yu, B. Lee, and H.-S. P. Wong, "Metal oxide resistive switching memory," in *Functional Metal Oxide Nanostructures*, pp. 303–335, Springer, 2012.

[31] J. J. Yang, M. D. Pickett, X. Li, D. A. A. Ohlberg, D. R. Stewart, and R. S. Williams, "Memristive switching mechanism for metal/oxide/metal nanodevices. Supp," *Nature nanotechnology*, vol. 3, no. 7, pp. 429–433, 2008.

[32] F. Pan, C. Chen, Z.-s. Wang, Y.-c. Yang, J. Yang, and F. Zeng, "Nonvolatile resistive switching memories-characteristics, mechanisms and challenges," *Progress in Natural Science: Materials International*, vol. 20, pp. 1–15, 2010.

[33] I. Valov, R. Waser, J. R. Jameson, and M. N. Kozicki, "Electrochemical metallization memories–fundamentals, applications, prospects.," *Nanotechnology*, vol. 22, no. 25, p. 254003, 2011.

[34] R. Waser, *Nanoelectronics and information technology*. John Wiley & Sons, 2012.

[35] A. Prakash, D. Jana, and S. Maikap, "TaOx-based resistive switching memories: prospective and challenges.," *Nanoscale research letters*, vol. 8, no. 1, p. 418, 2013.

[36] S. H. Chang, J. S. Lee, S. C. Chae, S. B. Lee, C. Liu, B. Kahng, D. W. Kim, and T. W. Noh, "Occurrence of both unipolar memory and threshold resistance switching in a NiO film," *Physical Review Letters*, vol. 102, no. 2, pp. 1–4, 2009.

[37] S. Yu, X. Guan, and H. Wong, "Conduction mechanism of TiN/HfOx/Pt resistive switching memory: A trap-assisted-tunneling model," *Applied Physics Letters*, vol. 99, p. 063507, 2011.

[38] J. J. Yang, M. X. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelley, G. Medeiros-Ribeiro, and R. S. Williams, "High switching endurance in TaOx memristive devices," *Applied Physics Letters*, vol. 97, no. 23, pp. 10–13, 2010.

[39] H. Jeon, J. Park, W. Jang, H. Kim, H. Song, H. Kim, H. Seo, and H. Jeon, "Resistive switching behaviors of Cu/TaOx/TiN device with combined oxygen vacancy/copper conductive filaments," *Current Applied Physics*, vol. 15, no. 9, pp. 1005–1009, 2015.

[40] C.-Y. Liu, J.-J. Huang, C.-H. Lai, and C.-H. Lin, "Influence of embedding Cu nanoparticles into a Cu/SiO2/Pt structure on its resistive switching.," *Nanoscale research letters*, vol. 8, no. 1, p. 156, 2013.

[41] H. Jeon, J. Park, W. Jang, H. Kim, C. Kang, H. Song, H. Kim, H. Seo, and H. Jeon, "Stabilized resistive switching behaviors of a Pt/TaOx/TiN RRAM under different oxygen contents," *physica status solidi (a)*, vol. 211, no. 9, pp. 2189–2194, 2014.

[42] Y. E. Syu, T. C. Chang, J. H. Lou, T. M. Tsai, K. C. Chang, M. J. Tsai, Y. L. Wang, M. Liu, and S. M. Sze, "Atomic-level quantized reaction of HfOx memristor," *Applied Physics Letters*, vol. 102, no. 17, 2013.

[43] L. Goux, A. Fantini, A. Redolfi, C. Y. Chen, F. F. Shi, R. Degraeve, Y. Y. Chen, T. Witters, G. Groeseneken, and M. Jurczak, "Role of the Ta scavenger electrode in the excellent switching control and reliability of a scalable low-current operated TiN/Ta2O5/Ta RRAM device," *Digest of Technical Papers - Symposium on VLSI Technology*, vol. 10, no. 2011, pp. 2013–2014, 2014.

[44] C.-Y. Liu, X.-J. Lin, H.-Y. Wang, and C.-H. Lai, "Improved resistive switching dispersion of NiOx thin film by Cu-doping method," *Japanese Journal of Applied Physics*, vol. 49, no. 5R, p. 056507, 2010.

[45] S. Long, L. Perniola, C. Cagli, J. Buckley, X. Lian, E. Miranda, F. Pan, M. Liu, and J. Suñé, "Voltage and power-controlled regimes in the progressive unipolar RESET transition of HfOx-based RRAM.," *Scientific reports*, vol. 3, p. 2929, 2013.

[46] J. S. Kwak, Y. H. Do, Y. C. Bae, H. Im, and J. P. Hong, "Reproducible unipolar resistive switching behaviors in the metal-deficient CoOx thin film," *Thin Solid Films*, vol. 518, no. 22, pp. 6437–6440, 2010.

[47] C. E. Shannon, "A symbolic analysis of relay and switching circuits," 1938.

[48] A. N. Whitehead and B. A. W. Russell, *Principia mathematica; 1st ed.* Cambridge: Cambridge Univ. Press, 1910.

[49] J. Borghetti, G. S. Snider, P. J. Kuekes, J. J. Yang, D. R. Stewart, and R. S. Williams, "'Memristive' switches enable 'stateful' logic operations via material implication.," *Nature*, vol. 464, no. 7290, pp. 873–876, 2010.

[50] E. Lehtonen, J. Poikonen, and M. Laiho, "Two memristors suffice to compute all Boolean functions," *Electronics Letters*, vol. 46, no. 3, p. 230, 2010.

[51] D. B. Strukov, D. R. Stewart, J. Borghetti, X. Li, M. Pickett, G. Medeiros Ribeiro, W. Robinett, G. Snider, J. P. Strachan, W. Wu, Q. Xia, J. J. Yang, and R. S. Williams, "Hybrid CMOS/memristor circuits," *ISCAS 2010 - 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems*, pp. 1967–1970, 2010.

[52] D. B. Strukov and K. K. Likharev, "CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices," *Nanotechnology*, vol. 16, no. 6, pp. 888–900, 2005.

[53] Y. V. Pershin and M. Di Ventra, "Practical approach to programmable analog circuits with memristors," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 8, pp. 1857–1864, 2010.

[54] T. Matsumoto, "A chaotic attractor from Chua's circuit," *IEEE Transactions on Circuits and Systems*, vol. 31, no. 12, pp. 1055–1058, 1984.

[55] M. Itoh and L. Chua, "Memristor oscillators," *International Journal of Bifurcation and Chaos*, vol. 18, no. 11, pp. 3183–3206, 2008.

[56] T. Driscoll, Y. V. Pershin, D. N. Basov, and M. Di Ventra, "Chaotic memristor," *Applied Physics A: Materials Science and Processing*, vol. 102, no. 4, pp. 885–889, 2011.

[57] P. Mazumder, S. M. Kang, and R. Waser, "Memristors: Devices, models, and applications," *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1911–1919, 2012.

[58] J. Sun, Y. Shen, Q. Yin, and C. Xu, "Compound synchronization of four memristor chaotic oscillator systems and secure communication," *Chaos*, vol. 23, no. 1, 2013.

[59] L. Chua and L. Yang, "Cellular Neural Networks: Theory," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 10, 1988.

[60] L. Chua and T. Roska, *Cellular Neural Networks and Visual Computing: Foundations and Applications*. New York, NY, USA: Cambridge University Press, 2002.

[61] D. Ielmini and R. Waser, eds., *Resistive switching — from fundamentals of nanoionic redox processes to memristive device applications*. Wiley–VCH, 2015.

[62] J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing.," *Nature nanotechnology*, vol. 8, no. 1, pp. 13–24, 2013.

[63] I. G. Baek, D. C. Kim, M. J. Lee, H. J. Kim, E. K. Yim, M. S. Lee, J. E. Lee, S. E. Ahn, S. Seo, J. H. Lee, J. C. Park, Y. K. Cha, S. O. Park, H. S. Kim, I. K. Yoo, U. Chung, J. T. Moon, and B. I. Ryu, "Multi-layer cross-point binary oxide resistive memory (OxRRAM) for post-NAND storage application," in *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest.*, pp. 750–753, 2005.

[64] A. Chen, "Accessibility of nano-crossbar arrays of resistive switching devices," *Proceedings of the IEEE Conference on Nanotechnology*, no. 3, pp. 1767–1771, 2011.

[65] J. Liang and H. P. Wong, "Cross-point memory array without cell selectors—device characteristics and data storage pattern dependencies," *Electron Devices, IEEE Transactions on*, vol. 57, no. 10, pp. 2531–2538, 2010.

[66] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 411–426, March 2007.

[67] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 30–42, Jan 2012.

[68] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[69] J. Misra and I. Saha, "Artificial neural networks in hardware: A survey of two decades of progress," *Neurocomputing*, vol. 74, no. 1-3, pp. 239–255, 2010.

[70] K. Seo, I. Kim, S. Jung, M. Jo, S. Park, J. Park, J. Shin, K. P. Biju, J. Kong, K. Lee, *et al.*, "Analog memory and spike-timing-dependent plasticity characteristics of a nanoscale titanium oxide bilayer resistive switching device," *Nanotechnology*, vol. 22, no. 25, p. 254023, 2011.

[71] Y. Wu, S. Yu, X. Guan, and H.-S. P. Wong, "Recent progress of resistive switching random access memory (RRAM)," *2012 IEEE Silicon Nanoelectronics Workshop (SNW)*, pp. 1–4, 2012.

[72] L. Zhao, H.-Y. Chen, S.-C. Wu, Z. Jiang, S. Yu, T.-H. Hou, H.-S. P. Wong, and Y. Nishi, "Multi-level control of conductive nano-filament evolution in HfO2 ReRAM by pulse-train operations.," *Nanoscale*, vol. 6, pp. 5698–702, 2014.

[73] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, "Training and operation of an integrated neuromorphic network based on metal-oxide memristors," *Nature*, vol. 521, no. 7550, pp. 61–64, 2015.

[74] F. C. Chiu, "A review on conduction mechanisms in dielectric films," *Advances in Materials Science and Engineering*, vol. 2014, 2014.

[75] C. Y. Chen, L. Goux, A. Fantini, A. Redolfi, S. Clima, R. Degraeve, Y. Y. Chen, G. Groeseneken, and M. Jurczak, "Understanding the impact of programming pulses and electrode materials on the endurance properties of scaled Ta2O5 RRAM cells," pp. 14.2.1–14.2.4, Dec 2014.

[76] R. Degraeve, A. Fantini, S. Clima, B. Govoreanu, L. Goux, Y. Y. Chen, D. J. Wouters, P. Roussel, G. S. Kar, G. Pourtois, S. Cosemans, J. A. Kittl, G. Groeseneken, M. Jurczak, and L. Altimime, "Dynamic 'hour glass' model for SET and RESET in HfO2 RRAM," *Digest of Technical Papers - Symposium on VLSI Technology*, no. 1, pp. 75–76, 2012.

[77] E. Miranda and J. Suñé, "Analytic modeling of leakage current through multiple breakdown paths in SiO2 films," *IEEE International Reliability Physics Symposium Proceedings*, vol. 2001-Janua, pp. 367–379, 2001.

[78] E. Miranda, C. Walczyk, C. Wenger, and T. Schroeder, "Model for the resistive switching effect in HfO2 MIM structures based on the transmission properties of

narrow constrictions," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 609–611, 2010.

[79] L. M. Prócel, L. Trojman, J. Moreno, F. Crupi, V. Maccaronio, R. Degraeve, L. Goux, and E. Simoen, "Experimental evidence of the quantum point contact theory in the conduction mechanism of bipolar HfO2-based resistive random access memories," *Journal of Applied Physics*, vol. 114, no. 7, 2013.

[80] X. Lian, S. Long, C. Cagli, J. Buckley, E. Miranda, M. Liu, and J. Suñé, "Quantum point contact model of filamentary conduction in resistive switching memories," *2012 13th International Conference on Ultimate Integration on Silicon, ULIS 2012*, vol. 1, no. 1, pp. 101–104, 2012.

[81] X. Zhu, W. Su, Y. Liu, B. Hu, L. Pan, W. Lu, J. Zhang, and R. W. Li, "Observation of conductance quantization in oxide-based resistive switching memory," *Advanced Materials*, vol. 24, no. 29, pp. 3941–3946, 2012.

[82] Y. Li, S. Long, Y. Liu, C. Hu, J. Teng, Q. Liu, H. Lv, J. Suñé, and M. Liu, "Conductance quantization in resistive random access memory.," *Nanoscale research letters*, vol. 10, no. 1, p. 420, 2015.

[83] T. Hiroi, "Evaluation of resistive random access memories operating under the current limitation by the use of a MOSFET," Master's thesis, Hokkaido University, 2014.

[84] A. Prakash, J. Park, J. Song, J. Woo, E. J. Cha, and H. Hwang, "Demonstration of low power 3-bit multilevel cell characteristics in a TaOx-Based RRAM by stack engineering," *IEEE Electron Device Letters*, vol. 36, no. 1, pp. 32–34, 2015.

[85] C. Turchetti, *Stochastic models of neural networks*, vol. 102. IOS Press, 2004.