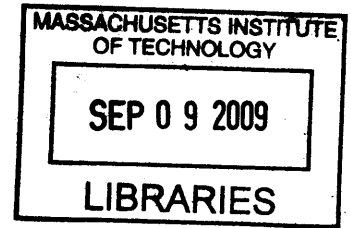# Evaluation of the Colossal Electroresistance (CER) Effect and Its Application in the Non-volatile Resistive Random Access Memory (RRAM)

by

Aulia Tegar Wicaksono

B. Eng. Materials Engineering
Nanyang Technological University, Singapore, 2008

SUBMITTED TO THE DEPARTMENT OF MATERIALS SCIENCE AND ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING IN MATERIALS SCIENCE AND ENGINEERING

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2009

Signature of Author: ....
Department of Materials Science and Engineering
July 30th, 2009

Certified by: ....
Caroline A. Ross
Toyota Professor of Materials Science and Engineering
Thesis Advisor

Accepted by: ....
Christine Ortiz
Chair, Departmental Committee on Graduate Students

Evaluation of the Colossal Electroresistance (CER) Effect and Its Application in the Non-volatile Resistive Random Access Memory (RRAM)

by

Aulia Tegar Wicaksono

Submitted to the Department of Materials Science and Engineering
on August 14th, 2009 in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Materials Science and Engineering

ABSTRACT

Flash memory, the current leading technology for non-volatile memory (NVM), is projected by many to run obsolete in the face of future miniaturization trend in the semiconductor devices due to some of its technical limitations. Several different technologies have been developed in attempt for replacing Flash memory as the most dominant NVM technology; none of which seems to indicate significant success at the moment. Among these technologies is RRAM (Resistive Random Access Memory), a novel type of memory technology which has only recently emerged to join the race. The underlying principle of an RRAM device is based on the colossal electroresistance (CER) effect, i.e. the resistance switching behavior upon application of voltage of varying polarity and/or magnitude.

This thesis aims to investigate the CER effect and how it can be designed to be a non-volatile memory as well as other novel application, e.g. memristor. The various technical aspects pertaining to this phenomenon, including the materials and the physical basis, are explored and analyzed. As a complementary to that, the market potential of the RRAM technology is also assessed. This includes the market study of memory industry, the current intellectual property (IP) landscape and some of the relevant business strategies. The production strategy (i.e. the production cost, initial investment, and pricing strategy) is then derived from the technical and market analysis evaluated earlier and with using some reasonable assumptions.

Thesis Advisor: Caroline A. Ross
Title: Toyota Professor of Materials Science and Engineering

Acknowledgments

Table of Contents

Chapter 6   Summary and future outlook

Appendices

References

List of Figures

List of Tables

List of Acronyms

1R1D	1-Resistor-1-Diode
1R1T	1-Resistor-1-Transistor
1R2T	1-Resistor-2-Transistor
ALD	Atomic layer deposition
BL	Bit Line
CBRAM	Conductive-bridging random access memory
CER	Colossal electroresistance
CVD	Chemical vapor deposition
DT	Domain-tunneling model
EPIC	Electric-pulse-induced capacitance
EPIR.	Electric-pulse-induced resistance
FE-RAM	Ferroelectric random access memory
HDD	Hard-disk drive
HRS	High-resistance state
IDP	Interface dipole polarization
LAO:YBCO	$LaAlO_3 : YBa_2Cu_3O_{7-x}$
LRS	Low-resistance state
MIM	Metal-insulator-metal
MOCVD	Metal-organic chemical vapor deposition
MRAM	Magnetic random access memory
NRAM	Nanotube random access memory
NVM	Non-volatile memory
PCM	Programmable metallization cell, or CBRAM
PCMO	$Pr_{1-x}Ca_xMnO_3$ ($x = 0.3$)
PCRAM	Phase change random access memory
RB-TS	Resistor-Bit (line) ‖ Transistor-Source (line)
RRAM	Resistive random access memory
RS-TB	Resistor-Source (line) ‖ Transistor-Bit (line)
RTA	Rapid thermal annealing
SCTO	$Sr_{1-x}Cr_xTiO_3$ ($x = 0.2$)
SECS	Strong electron correlation system
SL	Source Line
SSD	Solid-state drive
TMO	Transition metal oxides
WL	Word Line

Chapter 1

Non-volatile memory: introduction and its current development

The device scaling scenario as is projected by Moore's Law and further developed to be the ITRS report[1] calls for intense miniaturization of the microelectronic devices and components. One of the technological driving factors behind this race is the advancement of lithography technique that demonstrates improvement over the years. Among the devices, whose working principle has been constantly modified to adapt to this trend, is the computer memory.

1.1   Computer memory

Different types of computer memories are commercially available nowadays and each of them serves essential and specific function in any computing systems. One certainly can classify these memories according to the common ground they may share. A generally accepted way is one that puts them into two main categories: volatile and non-volatile memory.

1.1.1   Volatile and non-volatile memory

Volatile memory, or simply "memory" in everyday usage, typically refers to a form of computer memory that temporarily stores the data and requires constant power supply for the data maintenance. There are two types of memories that currently function as volatile memory: DRAM (Dynamic Random Access Memory) and SRAM (Static Random Access Memory). Their difference mainly lies on the fact that DRAM requires constant reading and rewriting the information with no modification (dynamic refresh) while SRAM does not[2].

In addition to that, DRAM has a simpler architecture that allows it to reach very high density and still cost cheaper while SRAM owes its complex structure for its faster operation time and easier interface control. Based on its characteristics, DRAM finds its application in high-capacity and low-cost devices such as the main memory in personal computers. On the other hand, SRAM operates mostly at devices where high reading and writing speed and low power consumption are of utmost consideration, such as automotive electronics, cell-phones, and cameras.

1

Non-volatile memory (NVM) is a type of memory that can retain the stored information even when the power has been plugged off. For this reason, this type of computer memory, commonly addressed as the "storage", is typically used for the task in which long-term persistent data storage is required. There are different types of storage technology serving as the non-volatile memory; each of which works under entirely different physical principles. Currently, the most predominantly used NVM is Flash memory; while several other technologies have been in competition aiming for the top positions in NVM industry. Flash memory is characterized as a high density memory with moderately fast reading access time. However, its slow writing mode along with the limited writing endurance renders it irrelevant for many applications. Further discussion on the issues that Flash memory predicted to encounter in the near future is presented in the next section.

1.1.2   Towards the universal memory

A universal memory element is one that exhibits scalability for high capacity of storage per unit volume, high read/write speed, compatibility with CMOS processes, as well as non-volatility. Such a memory technology would eliminate the need for multiple memories in many applications, improve the system performance and reliability by avoiding data transfer between multiple memories, and reduce overall system cost. Designing a novel type of memory that could satisfy almost, if not all, of the above requirements has been the ultimate goal of the research in this field. Many publications have reported the successful attempt of winning this pursuit. The actual results however seem to bring about some degree of questionability to such claim.

Several factors may contribute to the nullification of an idea that a particular type of memory can solely become the solution for the universal memory. Among the most fundamental ones are the economies of scale of the production, the possible slowdown in IC scaling, and the consumer's resistance behavior to technological revolution. Considering these factors, discussion on the universal memory would only be superficially attempted in this project.

1.1.3   Limits of the Flash memory[3]

Despite its current dominance in the nonvolatile memory industry, Flash memory may face some technical limitations in its future application that may render it obsolete; one of which is the

fundamental scaling limits due to the minimum oxide thickness required to sustain long-term data retention. Other issues that the research in memory technology attempt to solve are high internal programming voltage, slow programming speed, limited endurance and the most fundamental one, the limit on the dimensional scaling.

1)   Programming Voltage

In Flash technology, the minimum programming voltage is dictated by the minimum thickness of the tunnel dielectric. An extremely thin dielectric will cause the Fowler-Nordheim charge tunneling which will further obstruct the device's ability to maintain long-term data retention.

2)   Programming Speed

The programming and erasing speeds are controlled by the dynamics of charge injection or removal from the floating gate. In a typical Flash memory, the fastest writing time is in the range of 10 μs to 2 ms and is relatively slow as compared to volatile memories such as SRAM and DRAM where the range is in the order of 0.1 μs or shorter.

3)   Endurance

Endurance of a memory cell is measured from the number of write/erase cycles it can sustain before failure due to physical breakdown can occur. In Flash memory, endurance is limited by intrinsic charging of the tunnel oxides and damage to the tunnel oxides due to repeated charge injection through them, and is currently in the range of 10^5 to 10^6 cycles. Neither of these breakdown mechanisms are observed at a typical non-destructive electrical testing at the beginning of the life of the memory.

4)   Scaling

Scaling projections for a Flash memory cell is limited by the physical phenomenon that occurs during the operation. The effective channel length is limited by hot-electron operation, punch-through voltage, and leakage issues, while the effective channel width is limited by read current requirements and isolation encroachment of the field area. Considering both active and inactive area, the physical memory cell size is expressed as $A_{eff} = kF^2$ where F is the technology node minimum feature size and k is the parameter that varies for different type of memory cells. Based on the industrial trend, the physical cell size of Flash memory is expected to be $10F^2$ in the latest ITRS report (2007).

1.2   Future of the non-volatile memory industry

As the industry moves toward the lateral feature size projected for 2016 (i.e. beyond 45 nm)[4], the conventional memory technologies, particularly Flash memory, are rapidly approaching their miniaturization limits. To answer to these issues, different types of novel non-volatile memory are being developed. Some have already witnessed their commercialization; some are still under the prototype development stage, while other might have just been finalizing their conceptualization. These memories include RRAM (the focus of this project), MRAM, FeRAM, PCRAM, CBRAM and NRAM[5]; their acronyms are listed in page x.

1.2.1   Viable alternatives: PCRAM, CBRAM and RRAM

Among the alternative technologies mentioned above, some (e.g. Fe-RAM) are predicted to experience some struggle in adjusting to further scaling scenario. Such prediction is derived from the fact that their working principle is based on charge storage which proves to pose increasingly difficult reliability in retaining sufficient electron in the shrinking cells[5]. Nonvolatile memory concept aimed at the horizon beyond 2013 is suggested to be one that is based on resistance change rather than charge storage. Technologies that satisfy this requirement are *PCRAM, CBRAM and RRAM* [5]. This report will focus on the assessment of RRAM's potentials from the technology and market perspective. To provide with a better understanding on the position of RRAM as compared to the other leading candidates, the following technological summary is presented along with the illustration as shown in Figure 1.
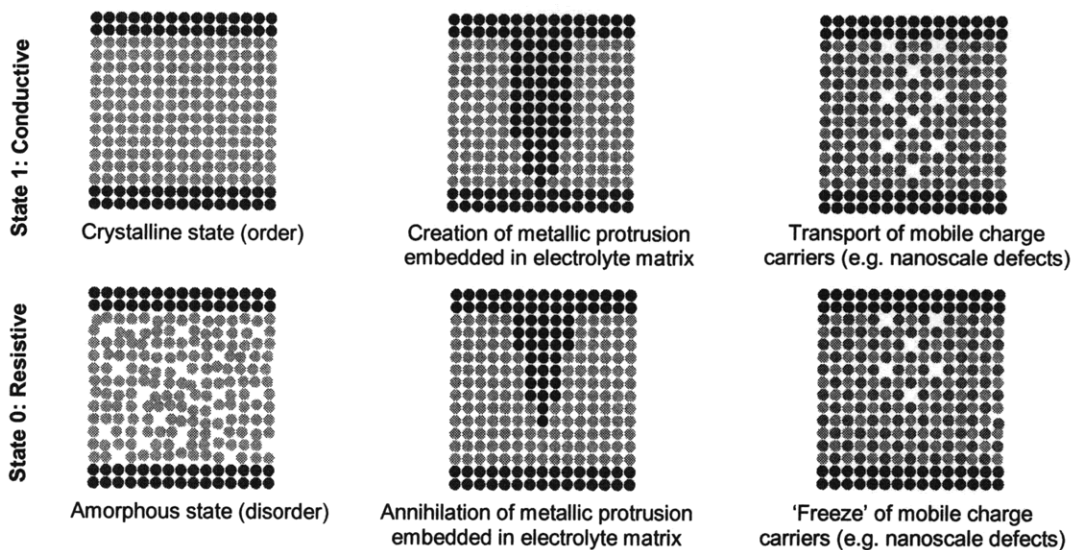
Figure 1 The working mechanisms of PCRAM (left), CBRAM (middle) and RRAM (right)[5]

1) PCRAM (Phase-Change RAM)

PCRAM employs the unique behavior of chalcogenide material. In a system where the chalcogenide is sandwiched between two electrodes (layer of black-dots in Figure 1), its resistance can be switched between two states—conducting crystalline state and resistive amorphous state [Figure 1 (top-left and bottom-left)]—upon heating. More detailed analysis of PC-RAM along with its market potential is presented in Ref. [6]

2) CBRAM (Conductive-bridging RAM)

CBRAM (also known as programmable metallization cell, PMC) is based on the physical relocation of ions within a solid electrolyte[7]. The memory cell is made of two solid metal electrodes separated by a thin film of electrolyte; one is relatively inert while the other is electrochemically active. When a negative bias is applied to the inert electrode, metal ions in the electrolyte, as well as some originating from the now-positive active electrode, flow in the electrolyte and are neutralized by electrons from the inert electrode. After a short period of time the ions flowing into the filament form a small metallic protrusion between the two electrodes [Figure 1 (top-middle)]. The metallic protrusion dramatically reduces the resistance along that path, which can be measured to indicate that the writing process is complete. This conductive bridge will annihilate upon the application of reverse potential [Figure 1 (bottom-middle)].

3) RRAM

The fundamental working principle of Resistive RAM remains a continuous debate among the scientific community; some of which are discussed in Chapter 3. Figure 1 illustrates one of the mechanisms for resistance switching in RRAM: the formation and rupture of conducting filament made of mobile charges during the applied potential. The layer of black-dots on the top and bottom surface represents each of the electrode. Originally, the RRAM prototype was designed on the basis of the hysteretic I-V curve found in the manganite perovskite films[8]. As more experimental results from different materials were reported, several other interesting properties have been identified, and collectively are called the electric-pulse-induced resistance change (EPIR) phenomenon or colossal electro-resistance. These properties further allow the design of a more robust RRAM cell capable of replacing Flash memory.

Chapter 2
Memory characteristics of the RRAM cell

A novel type of non-volatile memory, RRAM (resistive random access memory) has emerged into attention recently as one of the candidates for future memory technology. In its simplest structure, an RRAM cell consists of an insulator film sandwiched by two metal layers (e.g. Ag, Pt, Al, Ti, etc). The insulator film can be made of either transition metal oxide (TMO) [e.g. binary oxides ($NiO^{[9]}$, $TiO_2^{[10]}$), complex perovskites ($Pr_{1-x}Ca_xMnO_3$ (PCMO)[8], $Sr_{1-x}Cr_xTiO_3$ (SCTO)[11], $La_2CuO_4^{[12]}$)] or organic material (e.g. amino-imidazole dicarbonitrile AIDCN[13], dichloro-dicyano benzoquinone DDQ[14]) in which the electron correlation effects usually play an important role[15]. As will be explained in the later section, the thickness of the insulator film should be in the order of nanometer since at this scale, the phase separation and tunneling phenomenon that are relevant for resistance switching can still be of significance. Schematic representation of the prototype RRAM cell is illustrated in Figure 2. Dimensions of such prototype cell may vary from one report to another but they are typically within tens to hundreds nanometer.
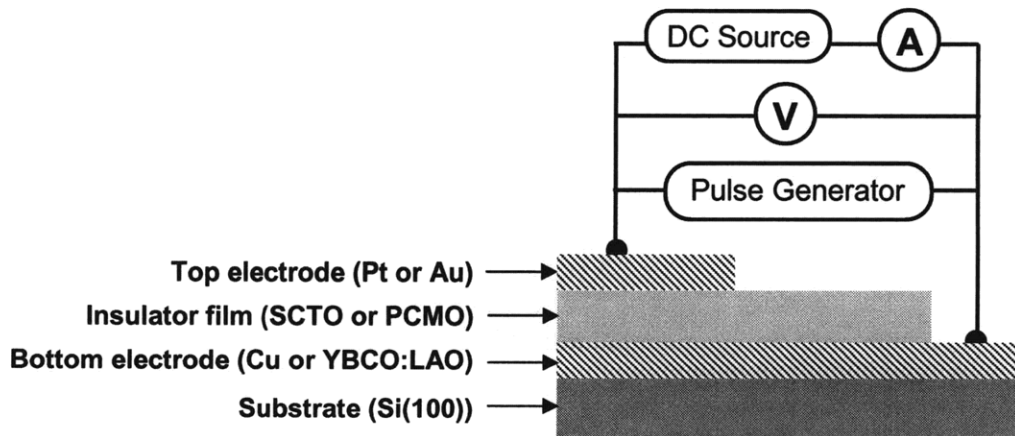
Figure 2 Schematic illustration of a prototype RRAM cell

The memory-related characteristic of the RRAM cell is observed when a voltage bias of varying magnitude and polarity is applied to the cell at room temperature and under the absence of external magnetic field. When the voltage sweep arrives at certain value, the resistance of the cell will dramatically change; the ratio between the high- and the low-resistance state (HRS and LRS) can be as high as two to three orders of magnitude and they can thus represent the '1' and

'0' state in binary logic. In addition to displaying two distinct accessible states, the RRAM cell must also satisfy other criteria before it can be further deemed as a potential candidate for non-volatile memory cell. A more elaborate discussion of these criteria is presented in the next sections. In doing so, this thesis will consider two of the most frequently investigated cells as the subjects of the study without losing the generality of the overall RRAM discussion: the Au/PCMO ($Pr_{0.7}Ca_{0.3}MnO_3$)/YBCO:LAO (or PCMO)[8] cell and the Pt/SCTO ($Sr_{0.8}Cr_{0.2}TiO_3$)/Cu (or SCTO) cell[11]. In addition to that, a recent study of binary oxide Pt/SnO_2/Pt cell[16] is also occasionally discussed. The experimental data is taken from several publications, primarily, by the Centre of Advanced Materials, University of Houston, TX and the IBM Research Division in Switzerland.

2.1 The existence of at least two distinct and easily accessible states

As discussed earlier, the simplest memory cell, a 1-bit cell, must exhibit two unambiguous states to store a '1' or '0'. In an RRAM device, these states correspond to the high- and low-resistance state of the MIM structure. The resistance switching is observed by measuring the current flowing through the device as a function of the applied voltage. In this measurement, the field is swept from a certain negative value (usually less than 10 V), increased at a small increment (typically about 0.25 V or less), to a positive value and then is reversed. The resulting current is plotted against the applied voltage and a hysteresis-like curve is obtained.

Since the resistance is inversely proportional to the slope of the curve, the stability of high- and low-resistance state is seen to be relevant only along certain voltage window where the ohmic (linear) behavior is observed. After the ohmic window has been identified, it will then be used as the operating voltage in future device application. In the case of PCMO and SCTO cells, their ohmic windows are ± 1.1 V and ± 0.8 V respectively. While the switching voltage often lies outside the operating window, their value is relatively small and still accessible for switching; i.e. twice the edge of the ohmic window for both PCMO and SCTO cells. In order to avoid the non-linearity behavior during the switching, the switching voltage is applied to the cell in the form of pulsed signal during the writing and erasing purpose. It will be shown later that the duration of such signal is in the nanosecond range. Furthermore, one theoretical model (the DT model,

7

Section 3.4) has particularly specified that such short switching signal is indeed necessary for some of the RRAM cells.
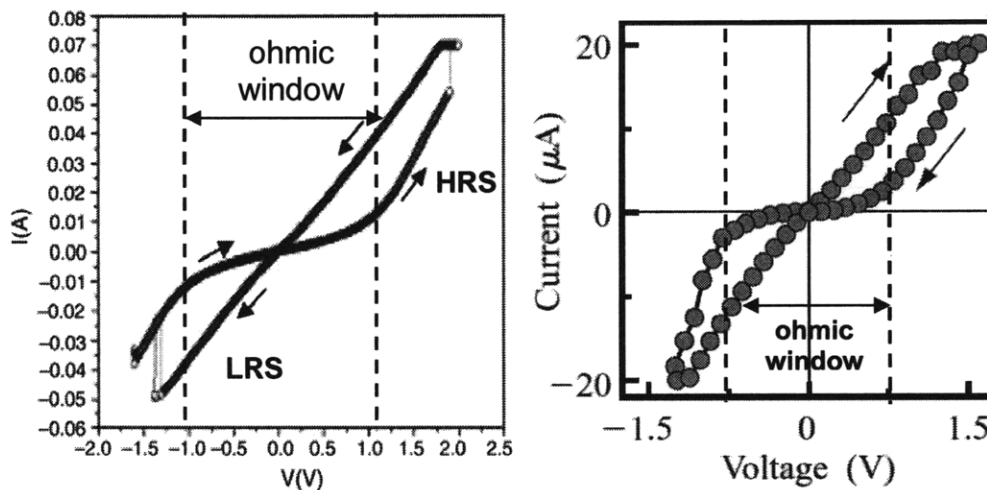
Figure 3 *I-V* hysteretic behavior of (left) PCMO cell[8], and (right) SCTO cell[11]

The symmetrical *I-V* characteristic of the PCMO and SCTO cells is not a common feature found in all known RRAM cells. Several works have reported the observation of unipolar switching in which the reversible switching occurs under the applied voltage of similar polarity but of different magnitude. A recent example of such system—most of which, interestingly enough, come from binary oxides cell—is Pt/SnO$_2$/Pt cell[16] whose *I-V* curve is drawn in semi-logarithmic scale in Figure 4.
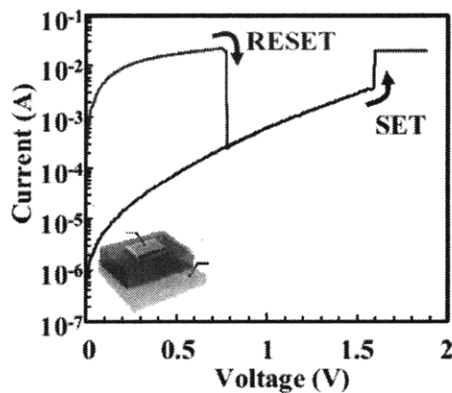
Figure 4 Unipolar switching observed in SnO$_2$ cell[16]

2.2   Switching time

The transition between the two resistance states is further studied by applying the set (ON-state, LRS) and reset (OFF-state, HRS) voltage consecutively. Experimental results showed that the

8

time it took for transition to occur in the prototype cell under investigation can be as fast as tens of nanoseconds (8 ns for PCMO and 130 ns for SCTO). Such ability to display rapid switching time is partially attributed to the nanometer-size of the prototype cell and, as will later be shown, this becomes greatly instrumental for the non-volatile memory application.
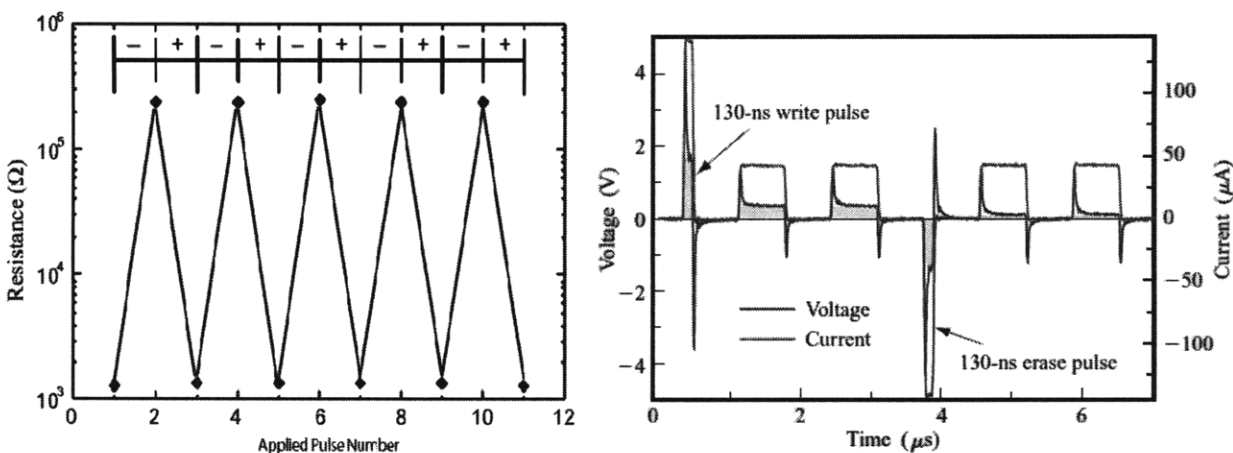
Figure 5 Transition time between the high- and low-resistance states; (left) PCMO cell[8] (input signal of 3.2 V applied within 8 ns), (right) SCTO cell[11] (input signal of 5 V applied within 130 ns)

2.3 Stability (non-volatility) of resistance state

It is of practical convenience to define the colossal electroresistance (CER) ratio* that illustrates the magnitude difference between the high- and the low-resistance state. Formal definition of this ratio is given as $CER \equiv \dfrac{\Delta R}{R} = \dfrac{R_{high} - R_{low}}{R_{low}}$. One of the characteristics of a good non-volatile memory cell is that it should be able to maintain the written data for considerably long period, e.g. 10 years. In RRAM cell, this translates to a condition where the CER ratio should be relatively constant for as long as the typical memory cell lasts. In addition to that, the stability of the CER ratio should also be maintained after certain number of set-reset cycles. The following plots illustrate the CER ratio stability over time and after certain number of cycles in SCTO (top figures) and tin oxide (SnO_2, bottom figures) cell. The extrapolation of the available data demonstrates the ability of the cell to retain the resistance state (both HRS and LRS) within the normal memory lifetime (e.g. 10 years or $\sim 10^8$ s). The data retention during voltage cycling, however, shows the fatigue behavior of the cell, particularly in the case of high-resistance state

---

* The origin of the colossal electroresistance (CER) terminology will be explained in Chapter 3

(HRS) whose value slowly decreases possibly due to the aggravation of the cell breakdown. Considering the HRS fluctuation during the set-reset cycle, it is then recommended to assume the most pessimistic value of the CER ratio, that is, to take the smallest possible R_{high} for setting the switching criterion, as is shown in the SnO_2 cell.
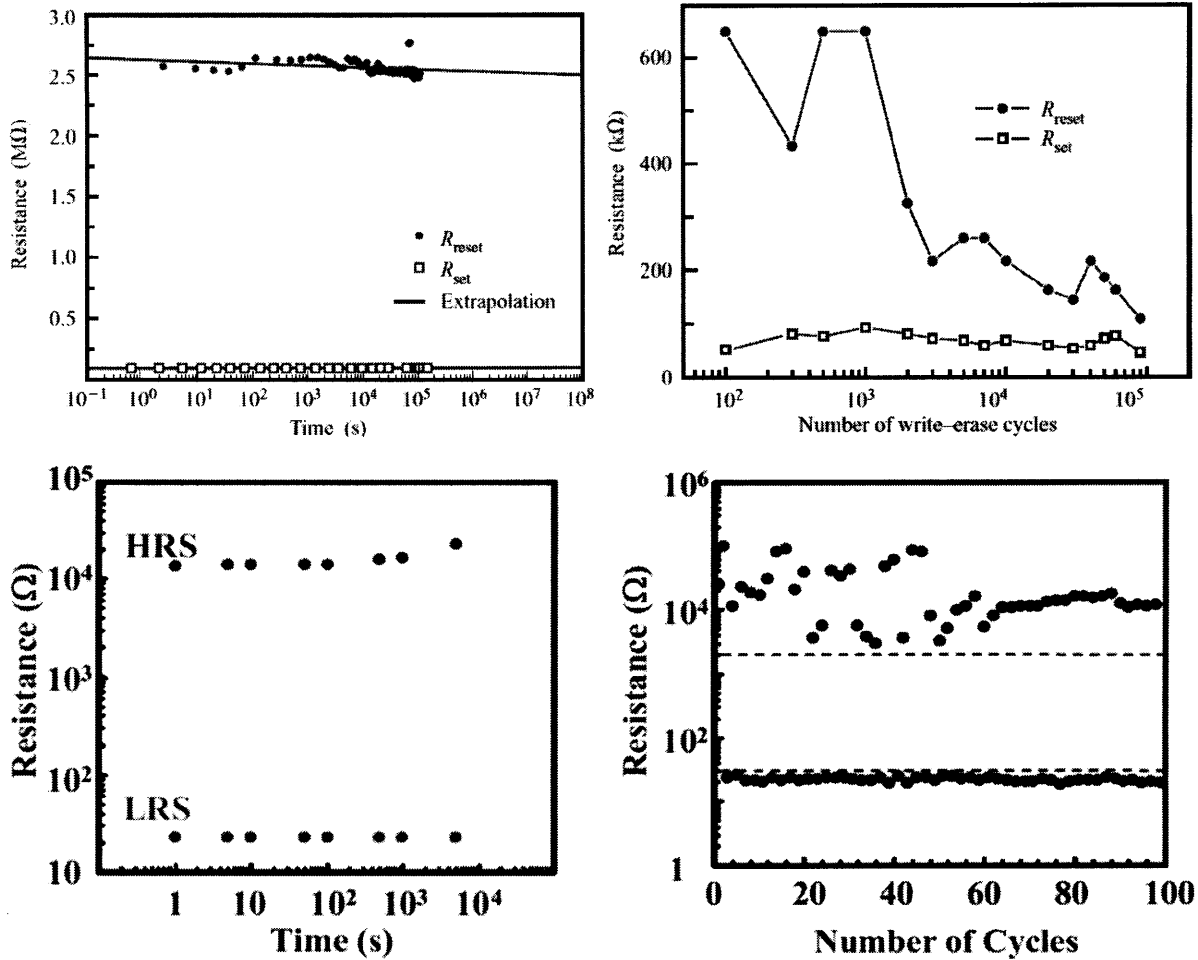
Figure 6 Stability of resistance states over time and upon set-reset cycles observed in: (top) SCTO cells[11], and (bottom) SnO_2 cells[16]

2.4   Multilevel states switching

Another striking property of RRAM cell is the observed multilevel switching. In the pulse operation (writing) mode, the actual value of low-resistance state (LRS) depends on the length and amplitude of the individual write pulses. Experimental results showed that by changing these parameters, different LRS levels can be addressed, each corresponding to a specific state of the memory. In this section, the data taken from the $Au/Sr_{1-x}Cr_xZrO_3$ (SCZO)/$SrRuO_3$[17] system is

10

presented. It is shown that there exist three different current levels, denoted as level 1 to 3 in Figure 7(b), corresponding to voltage pulses of different amplitude (6.5V, 5.3V, and 4.8V). The existence of four distinct levels (including level 0) demonstrates that two-bits of information can be stored in such simple structure. Currently, there are no known experimental evidences (let alone theoretical condition) that impose restriction to how many resistance states are accessible; thus, in principle, more levels could be addressed.

It should be noted that even though the results presented here were taken at 77 K, the same reference reported that the multilevel-switching behavior was also observed at higher temperature, e.g. room temperature at 300 K. The similar multilevel-switching behavior was also reported in PCMO and SCTO system; nonetheless, the omission of their discussion is taken since the published data[18,19] is unfortunately unclearly illustrated.
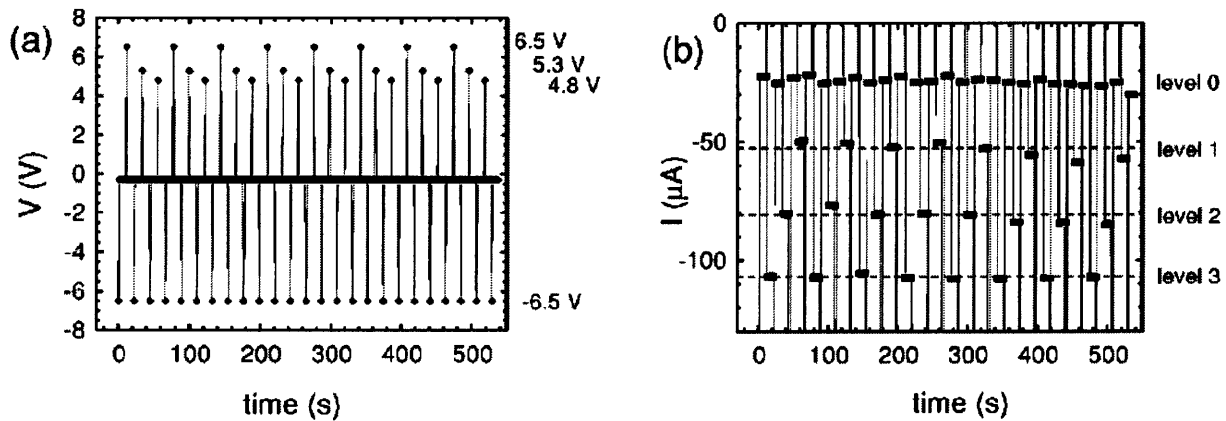
Figure 7 Multilevel switching in $Sr_{1-x}Cr_xZrO_3$-based RRAM cell operated at 77 K: (a) applied voltage input, (b) output current[17]

Chapter 3
Physics of colossal electroresistance (CER) effect

The resistance switching phenomenon in an M-I-M structure is investigated in many works under different names: the reproducible effect of resistive switching[17], electron instability effect[20], field-induced resistive switching[21], giant resistive switching[22], colossal electro-resistance[23], the electric-pulse-induced resistive (EPIR) switching effect[8].

Among these names, CER and EPIR are two of the most consistently adopted terminologies found in the various works investigating RRAM and its behavior. Throughout the discussion, this thesis will adopt the name CER with occasional mention of EPIR. The term colossal electroresistance itself was introduced in Ref. [23] where the resistance switching was observed in doped-manganites and named so using the analogy from the term colossal magnetoresistance (CMR)—one that is used to designate a change in the electric resistance in a magnetic field.

Since it was brought to the attention of scientific community, there have been several attempts to provide physical basis explaining the origin of the CER effect. This section discusses some of the notable theories along with their merits and weaknesses.

3.1   Polaron order-disorder transition model[24]

The CER phenomenon in manganites perovskite is considered. A strongly correlated material, doped-manganites $Re_{1-x}A_xMnO_3$ (Re: rare-earth and A: alkaline-earth cations) exhibit a variety of electronic and lattice interactions of equal strength and are thus able to create fundamentally differing electronic ground states. Among various interactions, two different basic types of electron-lattice coupling play a distinct role in the manganites:

(1)   Static crystal structure which undergoes a distortion due to Jahn-Teller effect[†,25] will generate internal stress on the Mn-O bonds whose effect might differ for the case of Re- or A-cations. The ramification of such stress is the lattice distortion that may in turn induce polar distortions or even the coupling of two or more different order parameters

---

† Jahn-Teller effect is observed when there is degeneracy in the lowest energy which prompts the crystal to split this energy level in order to remove the degeneracy and causing lattice distortion.

(e.g. multiferroicity). The extent to which this static coupling takes place is considered global, i.e. it is experienced throughout the bulk crystal. For this reason, this interaction is referred to as a long-range ordering.

(2)   The electron charge may in turn induce dynamic lattice distortion of the crystal that only takes place during the electron transport through crystal lattice. Such charge carrier can be pictured as being dressed by a cloth of lattice vibrations (phonons). This electron-phonon system can be represented by a quasi-particle excitation, namely a polaron. The length-scale at which this dynamic effect applies is considered local or short-range order. For that reason, the state at which this local dynamic effect becomes ubiquitous is called polaron liquid, while the case in which the previous global static effect is more prominent is called polaron solid.

Jooss et al[26] proposed that the field-induced resistance switching, i.e. CER effect, observed in many RRAM systems are attributed to the transition between these two 'phases': polaron order (solid) and disorder (liquid). The model is also supported by experimental results in which the TEM images show the phase transition that occurs *in-situ* during the application of electric field between the imaging tip and the sample.
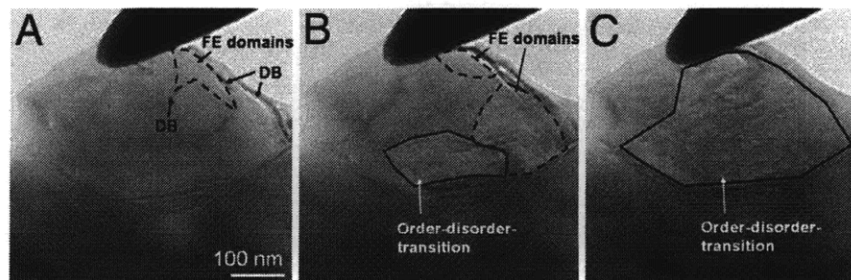
Figure 8 Low-magnification images of the PCMO grain showing electric domain patterns and the order-disorder transition at increasing applied current: (a) 0 μA, (b) 6 μA, and (c) 10 μA. DB and FE are domain boundary and ferro-electric, respectively[24]

Weakness:

Despite the merits and the depth of explanation which this model can offer, it lacks the universality for its application, especially in other RRAM materials that does not exhibit Jahn-Teller distortion. Nevertheless, the theoretical approach taken by this model can potentially be extended to include the resistance switching behavior of other materials.

3.2   Interface dipole polarization (IDP) model[27]

An initial applied dc field across the SECS (strong electron correlated systems) material could move charged defects, e.g. oxygen vacancies, causing dipole piled-ups or depletion at the metal-insulator interfaces, collectively known as the polarized interfacial layer or interface dipole polarization (IDP)[28]. An internal field whose direction is opposite with that of the external field is then built up as the result of this interface dipole polarization. When this external field is continuously applied, dipole may jump from one state to another. If the pulse period of this field fits the dipole relaxation time and its magnitude is greater than the threshold value, the dipoles will move between two metastable states that are represented as dual but unequal energy wells. The assumption of the unequal metastable energy states is derived from the experimental result in which switching endurance problem is often observed. Nonetheless, there have been other materials systems in which the endurance of the resistance states are permanent. In such systems, the energy of the metastable states is expected to be on equal level, allowing for non-volatile storage information.

Weakness:

Despite the thermodynamic approach taken by it, the IDP model still avoids many important aspects of the colossal electroresistance (CER). To name a few: the strongly correlated electron system that appears to be one of the most commonly shared characteristics of RRAM system has been ignored when discussing the resistance switching. While such point might be trivial, it has been shown by other models that this correlated electron effect is indeed one that gives rise to the CER phenomenon.

Other aspect that has escaped the IDP model description is the hysteresis characteristic of *I-V* curve. The fact that the ratio between high- and low-resistance state can reach as large as three to four order of magnitude and that the switching is reversible under the same absolute voltage indicate that these states cannot be simply explained only by the dipoles jump between the meta-stable state since this means that the transition must occur globally, i.e. throughout the metal-insulator interface. This would certainly require relatively high bias to overcome the potential barrier between two meta-stable wells which, of course, is not what has been observed.

14

3.3   The modified Schottky-barrier model[29]

The rectifying metal-insulator interface displays a resemblance with that of the Schottky contact. In addition to that, the condition of interface-state-induced band bending at the metal-insulator film interface is also assumed under the argument that when a density of interface states is high, electronic band in a semiconductor bends at an interface, independent with the work function of the metal electrode[30]. In conclusion, the model would assume that the degree of band-bending, i.e. barrier width and height, will depend on the net charge in the interface states and an energetical distribution of those in the band gap, as well as the band-bending, will cause rectification. The source of high interface states is assumed to be the charged defects, e.g. oxygen vacancies, typically found in TMO materials.
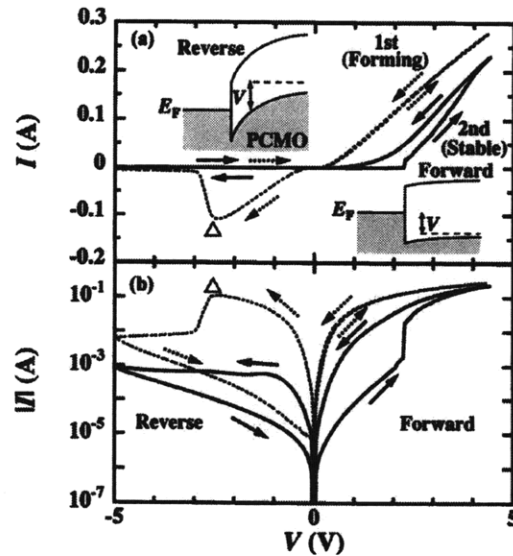
Figure 9 *I-V* characteristics of PCMO cell drawn in (a) linear and (b) logarithmic scale. Insets show electronic band diagram for Top-electrode/PCMO interface, providing support for the modified-Schottky barrier model[29]

On the basis of the interface-state-induced band-bending picture, Sawa et al[29] propose a model for resistance switching. By applying a sufficiently high voltage at the metal-insulator interface, a large amount of electrons is accumulated (extracted) into (from) the interface state upon reverse (forward) bias. Accordingly, a variation of a net charge in the interface states leads to a modification of a Schottky-like barrier width and/or height, because the degree of the band bending depends on a net charge in the interface states. This Schottky barrier modification during a complete voltage cycle is observed as the switching between resistance states as is schematically shown in Figure 9.

15

Weakness:

The universality of the model is questionable since the CER effect has been observed in many M_1-I-M_2 systems where M_1 and M_2 are two different metals, hence having different Schottky barrier. It also escapes to address the unipolar switching as well as the non-volatility feature of the system in which the resistance state is maintained even after the applied voltage is removed.

3.4 Domain-Tunneling (DT) or semi-phenomenological model

Most models attempt to describe the phenomenon from the microscopic perspective. While their meritorious aspects are acknowledged, these models lack the feature that fits the experimental result. Quantitative feature is instrumental to evaluate the validity of these views by comparing them with the macroscopic (empirical) observation. One model has attempted to introduce quantitative element into its description: the domain-tunneling (DT) model. In principal, this model considers the hysteretic behavior observed in the current-voltage (I-V) characteristic of different RRAM systems as the result of two physical mechanisms affecting the charge transport: the charge trap in the bulk and the change in the resistance at the metal-insulator interface[31].

3.4.1 Preliminary assumptions

Some features of this model are derived using the assumption of the insulating medium displaying the strongly correlated characteristics such as the doping-induced metal-insulator transition[32] and the spatial inhomogeneity that occurs at the nanoscale[33]. This insulating medium is described as non-percolating charged domains, corresponding to charge traps (e.g. dopants, vacancies, metallic clusters, nanodomains[9]) found in the real system. Upon the contact with the metal electrode, the insulating medium is being injected with a significant amount of carriers via electroforming process, generating 'conduction path' for current-tunneling. The amount of injected carriers however is insufficient to dope whole of domains. Only domains residing at or near the interface would be affected by these charge carriers—modifying their doping state relative to that of domains in the bulk. Thus, it can be approximated that there are only three types of domain: the top, the central and the bottom domain. These processes are illustrated schematically in Figure 10 [34].
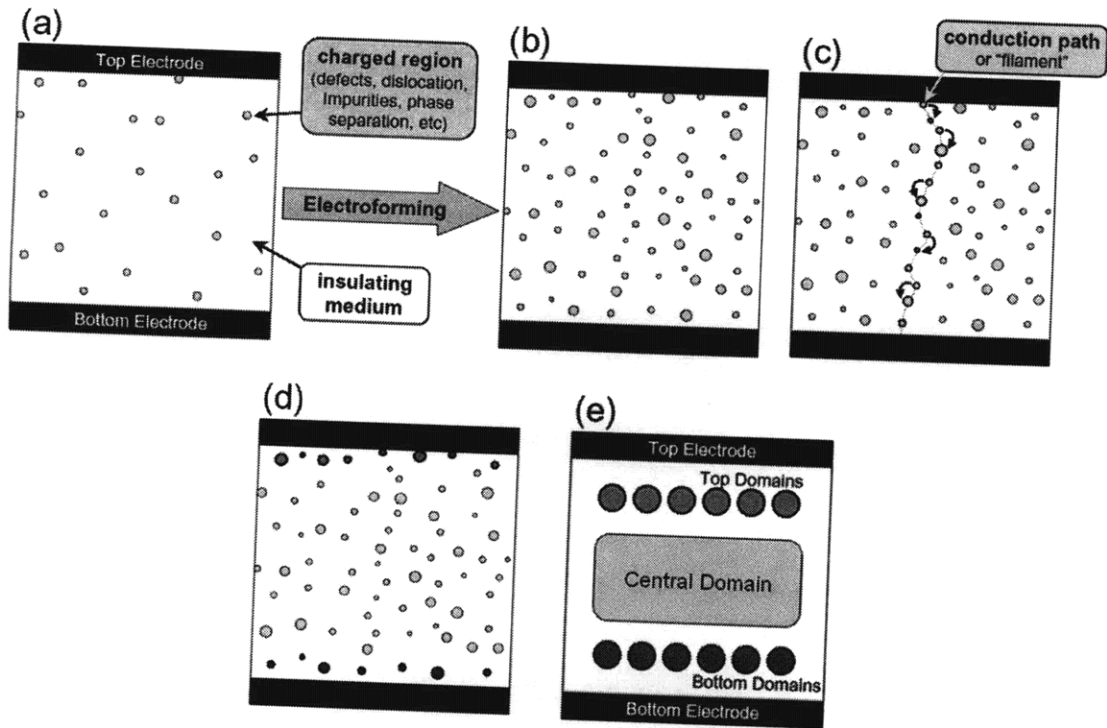
Figure 10 Schematic view of the processes that occur in the M-I-M memory cell leading to a structure with two electrodes sandwiching the insulating medium that contains smaller top-bottom domains and large middle domain[34]

3.4.2 Formulation of the model

The carriers tunnel between domains and electrodes under the action of an external field. The probability of charge transfer depends on phenomenological parameters such as the tunneling rates and the fraction of occupancy; all of which are combined in the charge transition probability equation that resembles Fermi's golden rule[35]:

$$I_{A \to B} = \frac{dn_B(t)}{dt} = \Gamma_{A \to B} \cdot N_A \cdot f_A(t) \cdot [1 - f_B(t)] \cdot g(V) \qquad \text{... Eq (1)}$$

where $I_{A \to B}$ is the current generated due to the charge transport from A to B, $\Gamma_{A \to B}$ is the tunneling rate, N_A is the total number of states in A, $n_A(t)$ is the number of carriers in A at time t, $f_A(t)$ is the fraction of occupancy of A at time t and defined as the ratio of the number of carriers occupying A and the total available states and $g(V)$ reflects the dependence of the transition probabilities on the applied external voltage V.

It is taken as convention that positive current transports carriers from the bottom electrode into the bottom domains and out of the top domains into the top electrode and its unit is the number

17

of carriers per unit of time. Using the above definition, the charge transport among the electrodes and the domains can be described by a set of three coupled nonlinear differential equations:

$$\frac{dn_B}{dt} = \left[\Gamma^{ext} \frac{N_E}{2}(1 - f_B) - \Gamma^{int} N_C f_B (1 - f_C) \right] \cdot g(V)$$

$$\frac{dn_C}{dt} = \left[\Gamma^{int} N_B f_B (1 - f_C) - \Gamma^{int} N_T f_C (1 - f_T) \right] \cdot g(V) \qquad \dots \text{Eq (2)}$$

$$\frac{dn_T}{dt} = \left[\Gamma^{int} N_C f_C (1 - f_T) - \Gamma^{ext} \frac{N_E}{2} f_T \right] \cdot g(V)$$

Several assumptions have been implicitly or explicitly included, namely:

(1)   The number of states are always half-filled

(2) $g(V)$ follows a hyperbolic function[36,37];

(3)   Tunneling rate only takes two forms: Γ^{int} for the tunneling between domains in the insulating material and Γ^{ext} for the electrode-domain tunneling ($\Gamma^{ext} \ll \Gamma^{int}$);

(4)   The interface tunneling rate depends on the charge accumulation, i.e. $\Gamma^{ext}_{T,B} = f(n_{B,T})$ [30], allowing the I-V hysteresis to cross at the origin, as is often experimentally observed.

(5)   Mott metal-insulator transition (mit) [21] (or any doping-driven transition[38]) in which the charge carriers become more localized as the applied potential is increased up to level at which the domains become close to half-filling, thus preventing current to be conducted (insulating)[22], and allowing the non-volatility characteristics. This effect is integrated via the following: $g_{mit}(V,t) = g_0(V) \cdot \exp(-\Delta / kT)$ under the condition that the domain occupation $f(t)$ is within a certain value, (e.g. 0.1) from half-filling[39]. Δ is the insulating energy gap opened during the transition to become insulator.

3.4.3   Results

Setting the parameters to values whose order of magnitude closely resembles the actual realistic numbers (or their approximation), the current-voltage behavior an RRAM system is modeled by using the Monte Carlo technique under the conditions that satisfy the DT model discussed earlier.

1.   Temporal progress of the applied voltage and the current

The external dc voltage bias is applied and set to follow certain protocol, as illustrated in Figure 11. The voltage bias mimics the bias in the typical memory cell application.

18
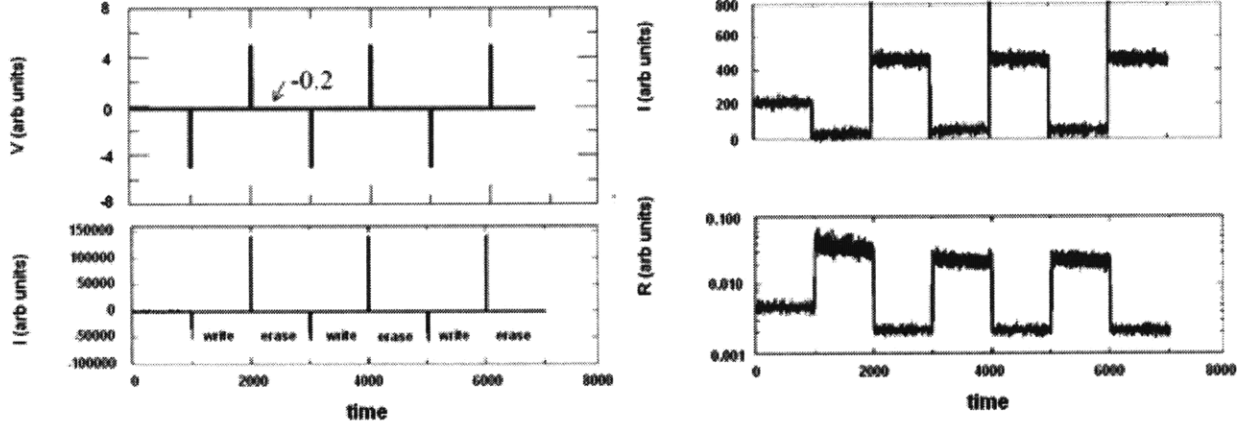
Figure 11 Temporal progress of the voltage, current and resistance in the RRAM modeled after DT-theory[31]

In the top panel, it is shown that an otherwise constant voltage is applied with some short and sharp pulses observed in certain points. These short pulses correspond to the write or erase voltage in the memory cell. The resulting current pattern (shown in the middle panel) follows the voltage trend with a distinct difference: Using the magnified scale, a more detailed current behavior is shown. Even though the voltage was set at a constant level after each pulse, the current behaves differently under a positive or negative pulse: high current after negative bias pulse reflects low resistance while low current after positive pulse reflects otherwise. Another interesting observation from this model is that the system is set to start at an arbitrary resistance. The action of a write pulse does not properly switch the system to "1" and only after an erase pulse is applied, the system gets properly initialized and begins switching between well-defined states. This behavior is not an artifact of the modeling, in fact, a variety of real RRAM systems observe the same characteristic[17].

2. Switching properties as indicated by the *I-V* hysteresis

The *I-V* relationship is obtained by setting the voltage protocol as $-V_{max} \to 0 \to V_{max} \to 0 \to -V_{max}$ and using as small voltage ramp as possible. The hysteretic behavior is modeled using the systems of equations (Eq. 2) and illustrated in Figure 12 (left).

19
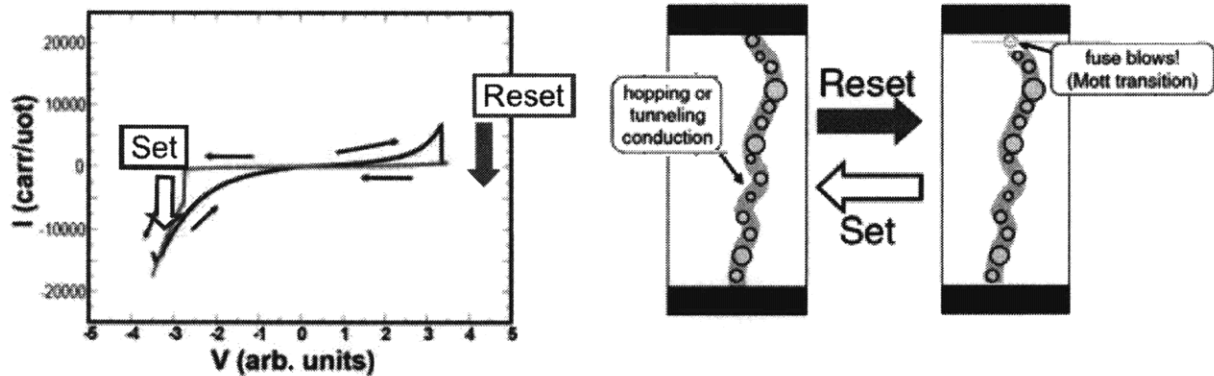
Figure 12: (left) Hysteretic *I-V* curve as plotted using the DT model; (right) resistance switching mechanism proposed by the DT model[34]

The physical mechanism underlying this hysteresis curve is explained as follows. Initially, the domains are assumed to be lightly filled, that is, in the metallic state. As the applied voltage is increased, charge is accumulated at the top domains, increasing their fraction of occupancy. Eventually, when they approach half-filling, they become insulating and open a Mott gap. The tunneling probability to and from the top domain is significantly decreased with consequent decline in the output current. With the reduced charge injection, the occupation level of the domain remains about constant and the current is low for the rest of the positive voltage sweep (Quadrant I). Such case is also applicable during the beginning of the negative voltage sweep in which the leakage of charge out of the top domain is also significantly suppressed, thus providing the nonvolatility feature of the high-resistance state. Eventually, a large negative *V* drives enough charge out of the top domains and the system backs to the LRS[37]. The schematic representation of the switching between two resistance states is shown in Figure 12 (right) where the reset current activates the Mott transition (HRS) while the opposite returns it back to conductive state (LRS).

3.4.4 Feature of unipolar switching

In most of MIM structure for RRAM cell, the nature of the switching is bipolar, that is, if a system is switched from one state to the other by a pulse of a given polarity, then a pulse of opposite polarity has to be applied to switch it back. In some cases, however, the resistance switching occurs by successive application of voltage pulses of the same polarity, as discussed in

Section 2.1. Such peculiarity was not included in the discussion on Section 3.4.2 earlier. In this part, a proposal for the model's extension to accommodate this behavior is presented[40].

Starting with the initial low resistance state (obtained via electroforming), the voltage is continuously applied until the bottom domain opens a Mott gap (half-filling condition), suppressing the tunneling current and activating the metal-insulator transition (HRS). To return it to the low-resistance state, voltage pulse of the same polarity, instead of reverse polarity, is applied. However, this pulse must satisfy certain condition, namely:

(1)   The magnitude should reach a threshould value V_{set} (~Δ) such that the Fermi level of the electrode aligns with the bottom of the empty conduction band of the bottom domain (BD) and a sudden injection of carriers occurs. When this condition is satisified, the exponential factor in $g_{mit}(V)$ vanishes and $g(V)$ now becomes $g_0(V)$ and the BD becomes metallized. The system is then back as the on-state and new cycle may begin.

(2)   This bias-induced transition should occur in a rapid fashion in order to prevent the BD from leaking the carriers down to the central domain. The duration for which the voltage pulse is applied to induce local breakdown at the BD should therefore be small compared to some characteristic transient time τ. The existence of parameter τ is experimentally confirmed, that is, the switching in voltage-sweep mode has to be performed under a rapid change of the applied bias[41,42].

Inclusion of the above assumptions does result in the feature of unipolar switching in the DT model[34]. Nevertheless, despite the qualitative agreement with the experimental result, the unipolar switching aspect is yet to be analytically integrated within the model, that is, a further modification to the system of non-linear differential equation in Eq. 2 is required.

3.5   Conclusion

Despite the variety of the CER models, some common grounds such as that presented by Waser et al[43] can be drawn. Before further proceeding, it is worth knowing that most of the published attempts did not present the following necessary information in details: the sample preparation, electrode deposition, the polarity of voltage applied during electroforming and during the first voltage pulse or voltage sweep, the electrod-pad size dependence of the current, the temperature

dependence of the electrical responses, the yield and the statistics in the cell-to-cell data, or the shift of characteristics on repeated cycling. These details are required for assessing the validity of the proposed explanation objectively.

3.5.1 Classification of the theoretical models

From the many theoretical models, there are generally three main dominantly contributing factors: a thermal effect, an electronic effect or an ionic effect[43].

A resistive switching based on a thermal effect typically shows a unipolar characteristic. It is originated from a potential-induced partial dielectric breakdown in which the material in a discharge filament is modified by Joule heating, and thus forms a weakly conductive path. During the reset transition, the conductive path is thermally disrupted by the high power density, making the path to rupture like a household fuse. The conducting filament theory proposed by Back et al[44] employs this mechanism for explaining the experimental result. Most of the prototype cells with binary oxides as the insulator behave according to this fuse-antifuse model.

Another basis for the colossal electroresistance is the electronic charge injection and/or displacement effect. There are at least two models that integrate this assumption in their description, namely the DT model and the modified Schottky barrier model. Despite the similarity, the origin for the electronic effect may differ from each model: the DT model assumes the Mott transition that is driven by the electronic charge injection observed in perovskite-type oxides, while the modified Schottky barrier model assumes that there exist charge-trapping sites (e.g. defects or metal nanoparticles) that affect the Schottky barrier at the interface.

The ionic transport effect is sometimes also attributed to the mechanism of CER. Here, it is proposed that electrochemical redox reactions provide the essential mechanism for bipolar resistive switching[45]. The continuously applied potential drives the mobile ions from the reduced/oxidized electrode. The drifted ions will be transformed into a highly conductive filament upon the discharge at the counterelectrode, resulting to LRS. When the polarity of the applied voltage is reversed, an electrochemical dissolution of the conductive filament takes place,

resetting the system into HRS. It should be noted that this ionic transport mechanism has been employed as a basis for other technology, that is, the conductive-bridging RAM (CBRAM).

Indeed, as discussed in Section 1.2.1, CBRAM and RRAM share an almost indistinguishable feature: both works via the utilization of the externally-induced resistance switching phenomenon. Nevertheless, a boundary line can still be drawn from both technologies: For CBRAM, metal ions dissolve in the material between the electrodes, requiring one electrode to be able to continuously provide the dissolving ions. In the case of RRAM, the material between the electrodes requires only one-time high applied electric field that causes local damage akin to dielectric breakdown—the electroforming step.

The classification of the theoretical models into three categories is itself inconclusive. Atomic-scale model such as the polaron order-disorder (Section 3.1), by definition, cannot be grouped into one of these categories despite the merits it has to offer. It should, however, also be noted that the reason for so might be that this model investigates different length-scale (less than 1 nm) compared to what other models do (100 nm – 1 μm). Noting the integration of Jahn-Teller effect, whose role is only significant in SECS material such as perovskite-type oxides, as well as the materials example accompanying the theory, one might suggest that the polaron order-disorder model and the DT model are two of the same model working on different length-scale.

3.5.2   Summary of the theoretical models

As a final note, the summary of the theoretical models is presented in the following table.

Table 1 Summary of the physical models proposed to explain CER effect

Theory	Proposer (Ref)	Category	Key Concept
Polaron order-disorder	Jooss et al[24]	Atomic-scale phenomenon (structure-related factor)	Transition between two 'phases' upon the reversible applied field: (1) Ordered state, i.e. global static coupling of two or more order parameters due to lattice distortion, and (2) Disordered state, i.e. dynamic coupling due to the local perturbation of the electron-phonon (polaron) system via the electron transport.

Interface Dipole polarization (IDP)	Pan et al[27]	Electronic (charge injection/ extraction)	Interface dipole jumps between several metastable energy states (corresponding to different resistance states). The level of these states may not be equal, resulting in a poor non-volatility of the memory cell.
Modified Schottky barrier	Sawa et al[29]	Electronic (charge injection/ extraction)	Electron accumulation/extraction in the interface state modifies the Schottky-like barrier of metal-insulator junction, resulting in different resistance states.
Domain-Tunneling (DT)	Rozenberg et al[31]	Electronic (charge injection/ extraction)	Charge trapping and detrapping occurs between the electrodes and the 'domains' (e.g. defects) of the insulator. Mott metal-insulator transition governs the resistance state of the system.
Formation/rupture of the conducting filament	Back et al[44]	Thermal disturbance	Field-induced dielectric breakdown in which the material in a discharge filament is modified by Joule heating, and thus forms a weakly conductive path. During the reset transition, the conductive path is thermally disrupted by the high power density, making the path to rupture like a household fuse.
Field-induced defects	Tsui et al[46]	Electronic (charge injection/ extraction)	Field-driven creation/annihilation of the interface defects determines the resistance states. The model resembles the DT model, except that it includes another assumption in which the defect states is dependent on the frequency (ω) of the applied field.
Oxygen diffusion	Ignatiev et al[47]	Ionic transport	Accumulation/depletion of oxygen vacancies at the interface regions modifies the resistance of the metal-insulator junction.
Electro-chemical metallization	Waser[45]	Ionic transport	Field-induced redox reaction activates the ionic conducting bridge (LRS) which is then dissolved upon the reversal of applied field (HRS).

24

Chapter 4

Application of CER effect: RRAM and other devices

Having discussed about the CER effect and the physics behind it, the thesis will now look at the technological application of this effect, particularly, in the non-volatile memory industry, i.e. the RRAM cell. In addition to that, some novel applications related to CER and the materials displaying it will also be briefly discussed in the later part of this chapter.

As is apparent from earlier chapters, the term 'RRAM' has been used interchangeably with 'CER effect'. Technically, RRAM is understood as the technological manifestation of CER in which the reversible resistance-switching of the MIM structure is used as the basis for storing information, i.e. HRS as '0' and LRS as '1'. An array of memory cells consisting of this MIM structure and other complementary devices, e.g. MOSFETs, is integrated into the CMOS circuit to form a working RRAM device. Several patents about RRAM architectures have been filed by and issued to various companies and research institutes since the first publication of CER effect (Section 5.2). These RRAM architectures are mostly derived from earlier memory technologies, such as PCRAM, and each of them has their respective merits and weaknesses. While the IP (patents) landscape related to RRAM will be analyzed later in Chapter 5, brief discussions about the memory architectures adopted from some of these patents would be essential, particularly for understanding the writing, reading and erasing operation of the memory cell. Before proceeding to the section about RRAM architecture, its materials and fabrications, a short summary of the RRAM is presented, including its characteristics as well as its feature comparison with other imminent memories: PCRAM and CBRAM (Section 1.2.1)

4.1   Characteristics

A simple prototype of an RRAM cell is illustrated in the following Figure 13. In this sketch, materials used to build the prototype cell resemble those of the structure discussed in Chapter 2 except for the bottom electrode in which Pt replaces LAO:YBCO electrode. This type of cell whose memory properties are summarized in the following list will from now on be used as the subject of RRAM analysis throughout the thesis.

1. Top electrode (Au)

2. PCMO film

3. Bottom electrode (Pt)

4. Substrate (Si)
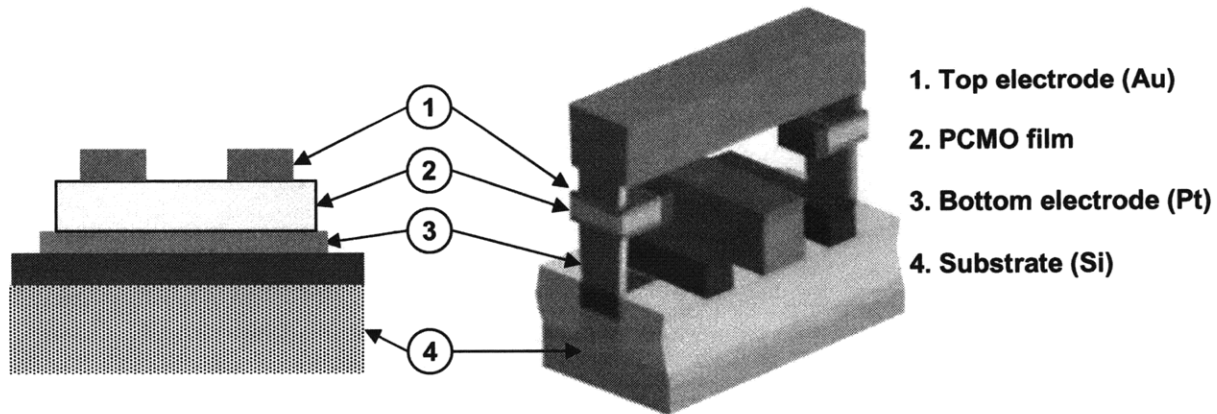
Figure 13 a prototype RRAM cell (left) and CMOS-integrated 1R1T memory cell (right)

Summary of the RRAM properties (from Chapter 2):

- Existence of at least two distinct resistance state. CER ratio can reach more than two orders of magnitude.

- Low switching voltage. The operating voltage window in this prototype cell is ±5V. Other materials are reported to operate from ±0.8V to ±9V.

- Good data retention ability. Within the experiment time interval (>12 months), the resistance state experience a slight and nonetheless tolerable change as the CER ratio is still within the detectable measure.

- Fast switching time. Transition between the resistance states can occur within 8 – 30 ns.

- Moderate endurance cycle. The minimum endurance cycle of an RRAM cell is shown to be inferior as compared to Flash memory, but still within the range of a typical non-volatile memory ($\sim 10^5$ cycles).

- Multilevel switching. Theoretically, there are no limitations to how many accessible resistance states are available in this material. The current experimental result demonstrates only four states, but may reach as high as six states.

4.2   RRAM comparison with other memory technology

In order to provide a much clearer understanding on how RRAM may compete in the non-volatile memory industry, the juxtaposition of RRAM properties with those of other memories is presented. This table shows that in most of the features, RRAM has considerably good advantage compared to other; exceptions being its endurance and the fact that it is still under embryonic

stage thus giving more rooms for further calibration which may require some adjustment to its features. Therefore, the perception in which RRAM is more superior, in terms of its characteristics, than the others is certainly unjustifiable despite the current data might suggest so.

Table 2 Feature comparison of the current and future memory technology[48,49,50]

Features	FLASH	PCRAM	CBRAM	RRAM
Evolution stage	Product	Adv. develop	Development (working prototype)	Embryonic development (prototype)
Smallest cell size	$5F^2$	$5.8F^2$	$4F^2$	$4F^2$ (1R1D), $6F^2$ (1R1T)
Scalability	≥ 22 nm	< 22 nm	5 nm	5 - 10 nm
Endurance	10^9-10^{12} cycles	10^8 cycles	$>10^6$ cycles	$> 10^6$ cycles
Write-erase time	1 µs/1-10 ms	10 ns/50 ns	35 ns/100 ns	10 ns/30 ns
Read time	10 µs	50 ns	50 ns	20 ns
Programming energy	65 pJ	100 pJ	1 pJ	1 pJ
Data retention	> 10 years	$\sim 10^2$ years	> 10 years	> 10 years
Multilevel bit operation	2 bit	3 bit	None	3-4 bit (max)

4.3 Architectures of RRAM device

This section will discuss two of the most common memory architecture—the resistor-transistor (RT) and the resistor-diode (RD) cell—along with their merits and drawbacks. Lateral view of these CMOS-integrated memory cells are sketched in some of the illustrations presented in the following discussions about each of the cell architectures. It is clear that in these layouts, the reversible resistance-switching MIM device is represented as variable resistor.

4.3.1 Resistor-Transistor (RT) memory cell

Three types of RT memories are discussed: 1R1T (RB-TS), 1R1T (RS-TB) and 1R2T.

Cell 1: 1R1T (Resistor-Bit Line || Transistor-Source Line, RB-TS)[51]

A memory cell is formed by connecting the variable resistor and the transistor in series, as illustrated in Figure 14; resistor connected to the Bit Line (BL) while transistor connected to the Source Line (SL). The layout of the complete circuit is shown in Figure 15. The single memory cell measures of $6F^2$.
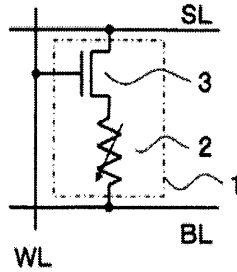
Figure 14 Single 1R1T (RB-TS) memory cell (1) with the resistor (2) connected to Bit-line (BL) and the transistor (3) connected to Source-line (SL)[52]
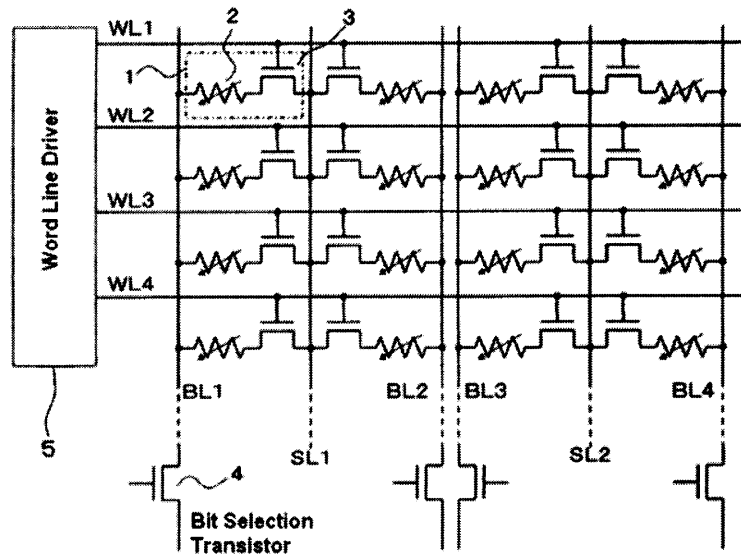
Figure 15 Circuit containing array of 1R1T cells; 1-3 (refer to Figure 14), 4 (bit-selection transistor) and 5 (the word-line driver)[52]

- *Reading operation*

The cell at the top right corner is the focus of the discussion. A sufficiently small bias for accessing the resistance state in the resistor (2) is applied to the Bit Line (BL1) by the Bit Selection Transistor (4). At the same time, the Word Line (WL1) which is connected to the gate of the Cell Selection Transistor (3) is set to a high potential (e.g. 7 V) by the Word Line Driver (5) in order to turn on this Cell Selection Transistor (3). Furthermore, by setting the source of the transistor (3) to a reference voltage (e.g. 0 V) via the Source Line (SL1), a current from the Bit Selection Transistor via BL1 to the resistor (2) and transistor (3) is generated. For the rest of the unselected cells, setting the level of unselected Word Lines and Bit Lines (WL2-4, BL2-4) to high impedance will cause the current route passing through other than the top rightmost cell to disappear. Under these

28

conditions, only change of resistance of the selected cell appears as a change of currents circulating via the Bit Line (BL1). By determining the current change using a reading circuit, it is possible to read the information stored in a selected cell.

- *Writing/Programming operation*

  A case in which the resistance of the variable resistor (2) is larger than a certain reference value when switched to the programming state and a case in which the resistance is smaller than the reference state when switched to the erasing state are assumed. In this case, the Bit Selection Transistor (4) is operated such that a programming voltage (e.g. 3 V) can be applied to the variable resistor (2) via the selected Bit Lines (BL1). The rest of the writing operation is similar to the reading operation: Cell Selection Transistor (3) is turned on by the Word Line Driver (5) via the Word Line (WL1). The Source Line (SL1) sets the source of Transistor (3) at ground potential. As a result, current is generated from the Bit Selection Transistor (4) to the memory cell activating the HRS and the data is programmed. Data will not be programmed in the unselected cells since all of the unselected Word Line (WL2-4) is set to a zero level, preventing the generation of current.

- *Erasing operation*

  There are two types of erasing operation: selective-erasing and block-erasing. Selected erasing is essentially the reverse of writing operation. Introducing opposite bias to the cell (i.e. high-potential in the SL and low-potential in the BL) would return the resistance to its original state, as is explained in the hysteretic *I-V* curve earlier. The block-erasing operation, on the other hand, involves simultaneous erasing of all data in every block. Every Bit Selection Transistors (4) is set to zero potential, making all resistors (2) to be in zero potential as well. At the same time, every Word Lines (WL1-4) is set to high potential (e.g. 7 V) turning on all Cell Selection Transistors (3). Moreover, all of the Source Lines (SL1-2) are then set to a reference voltage (e.g. 3 V). As a result, current is generated from the Source Line (high potential) via the Cell Selection Transistor (3) and variable resistor (2) to the Word Line (zero potential). This occurs in all memory cells across the circuit, and thus erasing operation is completed.

- *Advantages and Drawbacks*

  In the above configuration, the selected resistor (2) is connected to the selected Bit Line (BL1) but so are those unselected resistors. Therefore, when applying a bias voltage via

the Bit Line (BL1)—such in case of the reading operation—the voltage stress may also be experienced by the unselected resistors even though the Word Line (WL2-4) in the unselected row is kept at low level potential. Even if the voltage stress is so weak that it can be ignored for a one-time reading operation, this stress may be repeatedly generated in the same memory cell. Therefore, resistance states of the RRAM device may be slowly changed for a long time. This voltage-stress cycling problem not only occurs during the reading but also in the programming/writing operation.

Cell 2: 1R1T (Resistor-Source Line || Transistor-Bit Line, RS-TB)[51]

A slight modification to the previous memory cell is proposed. Here, the resistor and the transistor switch their places, as illustrated in Figure 16 while the layout of the complete circuit is shown in Figure 17. As the resistor is now electrically isolated from the Bit Line, the voltage stress problem is expected to be solved. The area of its single memory cell is similar, about $6F^2$.
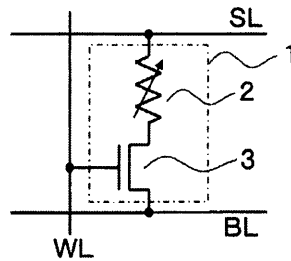
Figure 16 Single 1R1T (RS-TB) memory cell (notations are similar to Figure 14)
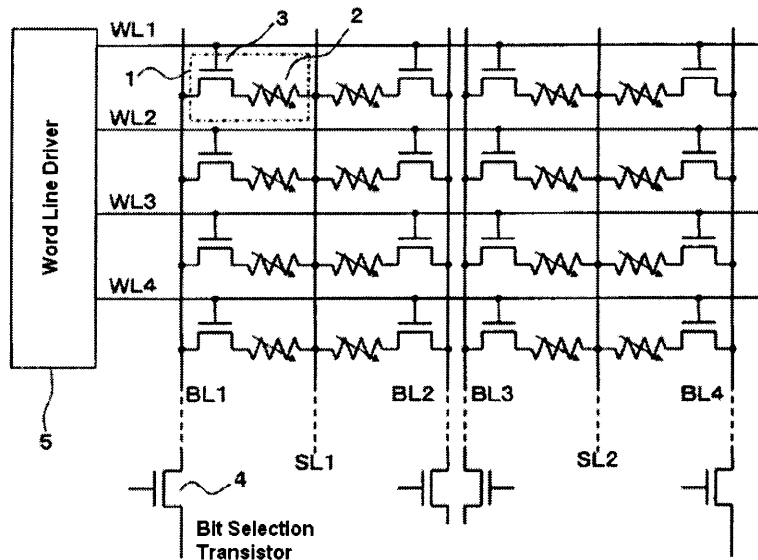
Figure 17 Circuit containing array of 1R1T cells; 1-3 (refer to Figure 16), 4-5 (as illustrated above)

30

- *Reading, Writing and Erasing*

  Similar reading and writing operation applies and therefore is succinctly explained here. A predetermined reading or writing bias (e.g. 1.5 or 3V) is applied to the selected cell via BL1 and at the same time, a WL connected to the gate of the transistor (3) is set to a high level (e.g. 7V) by the WL Driver (5), turning on the transistor (3). Moreover, by setting the SL1 at ground potential, current will flow from BL1 to SL1. In the case of reading, the change in resistance is recorded while in the case of writing, the change of resistance is generated. For the unselected cells, WL2-4 are grounded, providing electrical isolation to them. As for erasing operation, both block-erasing and selective-erasing are possible.

- *Advantages and Drawbacks*

  Since the resistor of unselected cells is isolated from the high-potential Bit Line (BL1), a voltage stress would not be developed in the unselected resistors even when the same BL (BL1) experience repetitive bias during the reading or writing. This eventually means that the possible electrical fatigue effect due to the voltage stress is greatly reduced, improving the reliability of the RRAM device. The issue in this architecture is somewhat related to that of the previous one, the 1R1T (RB-TS) cell. In this case, the direct connection of resistor to the SL is one that might pose problem during the erasing operation—the disturb phenomenon. This phenomenon refers to the unintentional erasing of the stored data and might occur because of the voltage stress that the unselected variable resistors must endure. During the selective-erasing operation, high potential is applied to the SL1, allowing the current to flow in the direction opposite with writing/reading operation, i.e. from SL1 to BL1. Even though the unselected Word Lines (WL2-4) is set at ground potential, turning off all the unselected transistors, voltage stress might still be experienced by every resistors connected to SL1 which may or may not contain stored data. As a result, resistance state of these unselected resistors may be returned to the original state, hence the unintentional erasing.

Cell 3: 1R2T (Transistor-Bit Line || Transistor-Source Line, TB-TS)[52]

In this configuration, the variable resistor is being 'guarded' by the transistor at both ends, thus preventing it from the voltage stress problem that might occur in the previous two configurations. The illustration of the single 1R2T memory cell whose measurement is $8F^2$ is shown below.

31
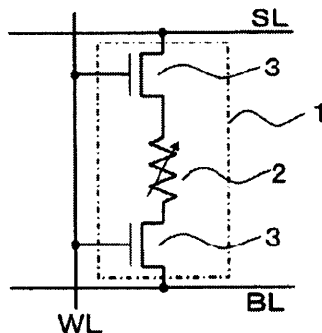
Figure 18 Single 1R2T memory cell (notation are similar to Figure 14)

- *Advantage and drawbacks*

The reading, writing and erasing operation are essentially the same with the 1R1T (RS-TB). The incorporation of another transistor helps to avoid the voltage-stress issue, though it may add some additional time in the switching time. Furthermore, the size of the cell increases significantly, reducing the device density.

4.3.2 Resistor-Diode (RD) memory cell[53,54]

Due to its rectifying nature, the RD design, particularly 1R1D, can only be applicable in a cell whose variable resistor is governed by the unipolar switching mechanism (Section 2.1). Despite the narrow range of materials selection—only few known CER materials observes unipolar switching—this architecture has a salient advantage in that it only measures at $4F^2$, making it the smallest configuration of RRAM memory. The most apparent contrast with RT-structure, apart from the fact that diode replaces transistor, is that there is no Source Line (SL) included. This would significantly modify the data operation (reading, etc) as will be explained shortly. The single 1R1D cell is illustrated in Figure 19 while the block cell is given in Figure 20.
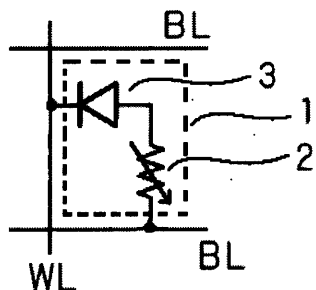
Figure 19 Single 1R1D memory cell (1) with the resistor (2) and diode (3) connected in series
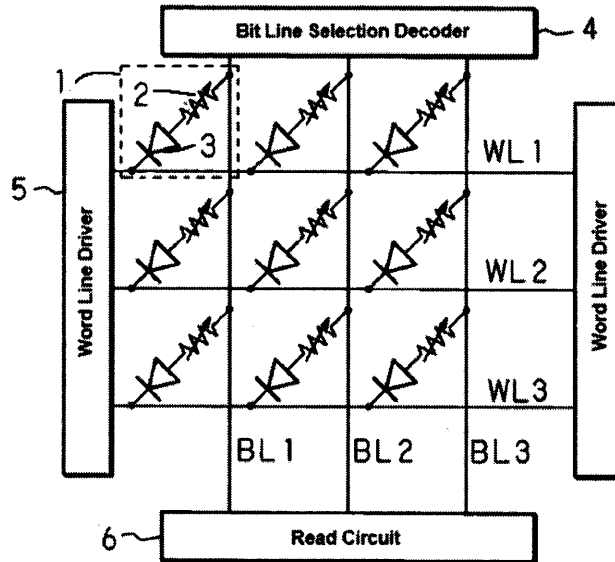
Figure 20 Circuit containing array of 1R1D cells; 1-3 (refer to Figure 19), 4-6 (as illustrated above)

- *Writing*

  The cell at the top right corner (dashed box) is the focus of the discussion. A writing voltage V_w and ground voltage are applied to the BL1 and the WL1 respectively. Since the Schottky diode (3) is forward biased, the resistance will change and data storage is performed. Other cell on the same column and not connected to the BL1 is called half-selected cell (BL-selection). A voltage of $V_w/2$ is applied to the word lines WL2 and WL3 so that the potential difference between both ends of the half-selected cell (BL-selection) may become $V_w/2$. The same also applied to the cells connected to WL1—half-selected cell (WL-selection). In their case, the voltage of $V_w/2$ is applied to BL2 and BL3 so that the potential difference between both ends of the half-selected cell (WL-selection) is also $V_w/2$. In other words, by setting the writing voltage V_w such that writing into the variable resistor (2) may not be performed when the potential difference between both ends of the half-selected cells is $V_w/2$, writing into these half-selected cells is prevented.

- *Reading and erasing*

  Selective-erasing operation works similarly with the writing operation. Here, the erasing voltage V_e applied to the BL1 follows the experimental result (such as shown in Figure 4) while ground potential is applied to WL1. Such operation will reset the resistance of resistor (2) to its original state. In the case of block erasing, the erasing voltage is applied in all Bit Lines so that each memory cell will return to its '0' state. The same writing and

selective-erasing mechanism is also applicable for reading operation. The only caveat in this operation is that the reading voltage must be carefully chosen such that it is distinct from other voltage operation. In a typical uni-polarly switched CER material, V_w or V_{set} is higher than V_e or V_{reset} (i.e. $V_w > V_e$). In that case, the reading voltage V_r can be set at value higher than V_w ($V_r > V_w$). If the opposite applies (i.e. $V_w < V_e$), then reading voltage V_r should be set at any value that is within the following range: $0 < V_r < V_w$.

- *Advantages and Drawbacks*

The most apparent advantage is certainly its size. Nevertheless, there are several drawbacks that may obstruct the use of this structure, e.g. the narrow selection materials choice and the voltage stress problem found in 1R1T cell. The various level of bias applied to Bit Line may also pose some problems in which the data operation (writing and erasing) might have some confusion if their respective voltage is not clearly distinct.

4.4 Fabrication method of the RRAM memory cell[52,53,54]

As are described in several patents, the production of an RRAM cell, regardless of its configuration, does not require highly specialized methods; they only involve typical microstructure fabrication techniques since its architecture is basically derived from the earlier CMOS-integrated memories. Furthermore, the fabrication techniques for some specific materials (e.g. PCMO) have also been available in the recent literatures. In this section, an elaborate discussion of complete RRAM fabrication is attempted. In order to help with the illustration, two architectures will be used: 1R1T and 1R1D.

4.4.1 Fabrication of 1R1T cell
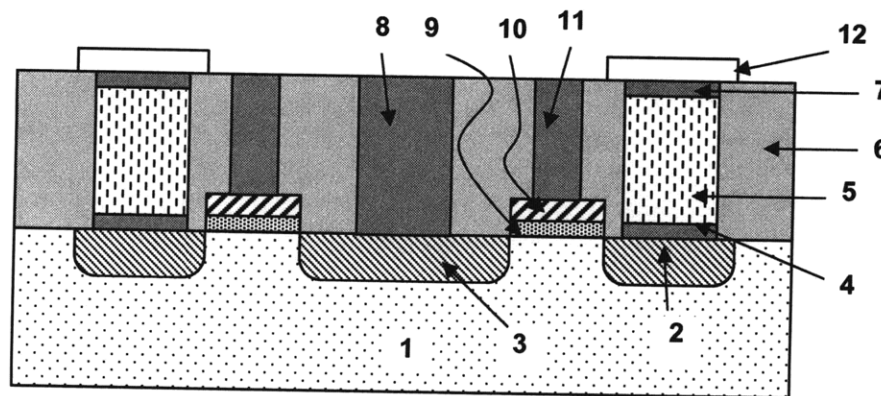
Legend:
1. Substrate (*p*-type)
2. N+ (conducting line)
3. N+ (conducting line)
4. Bottom electrode (Pt)
5. PCMO (resistive component)
6. Dielectric layer (SiO_2)

7. Top electrode (Pt)
8. Cu (plurality of contact window, Source Line)
9. Dielectric (ZrO_2 or HfO_2)
10. Metal (TiN, together with 9 to form gate structure, plurality of Word lines)
11. Cu (plurality of contact window, Source Line)
12. Cu (Bit Line)

Figure 21 Schematic lateral view of 1R1T cell along with its material components as proposed in [52]

A *p*-type silicon substrate (1) is provided. N^+-Si conducting line (2, 3) is introduced to the substrate using ion implantation to improve the transport along the junction between semiconductor substrate (1) and the subsequent metallization (8) or the bottom electrode (4). Dielectric layer (e.g. silicon dioxide) acting as insulator (6) is deposited on top of the substrate using thermal oxidation under dry ambient or sputtering. Using optical or electron-beam lithography (depending on the desired feature dimension), trenches for transistor's gate (9, 10), interconnect as well as the variable resistor (4, 5, 7) are formed. The transistor is first assembled by depositing the high-*k* dielectric layer, such as ZrO_2[55] and HfO_2[56] using atomic layer deposition (ALD) or MOCVD. The use of these materials is driven by the increasing need for further miniaturization (beyond 45nm) without the concomitant leakage (tunnelling) current found in silicon dioxide. In addition to high-*k* gate dielectric, metal gate is also employed to replace polysilicon currently not ideal due to its high electrical resistance. The choice of materials includes TaN and TiN[57] which is deposited using CVD or ALD[58] and sputtering[59] respectively. Metallization such as contact window above the gate (9,10) is formed using copper. Due to the high-etching resistance, copper must be fabricated via eletrodeposition process.

The resistive component (4,5,7) is first formed by Pt electrode deposited using ALD[60] or sputtering, continued with PCMO via several techniques, and finally the top Pt electrode. Before depositing the metal electrode, a thin layer of diffusion barrier such as TaN (not shown) is deposited. There are many methods used for PCMO thin film deposition, for example, PVD, MOCVD, sputtering[61] and spin-coating. Of these methods, spin-coating[62] is considered as one of the best candidates for large-scale manufacturing as it is a relatively low cost and easy process and thus will be used in the rest of the discussion. Despite its simple implementation, the spin-coat deposition discussed in the succeeding diagram may also poses some drawbacks; the most

apparent one is the baking and annealing which take place at high temperature and thus might cause significant degradation to other components. A low-temperature spin-coat deposition has been proposed in other patent[63], however, such method is currently applicable only for particular substrate, i.e. iridium substrate. In subsequent discussion, it is assumed that spin-coating deposition as described by Figure 22 will be implemented and possible degradation of other components is negligible.
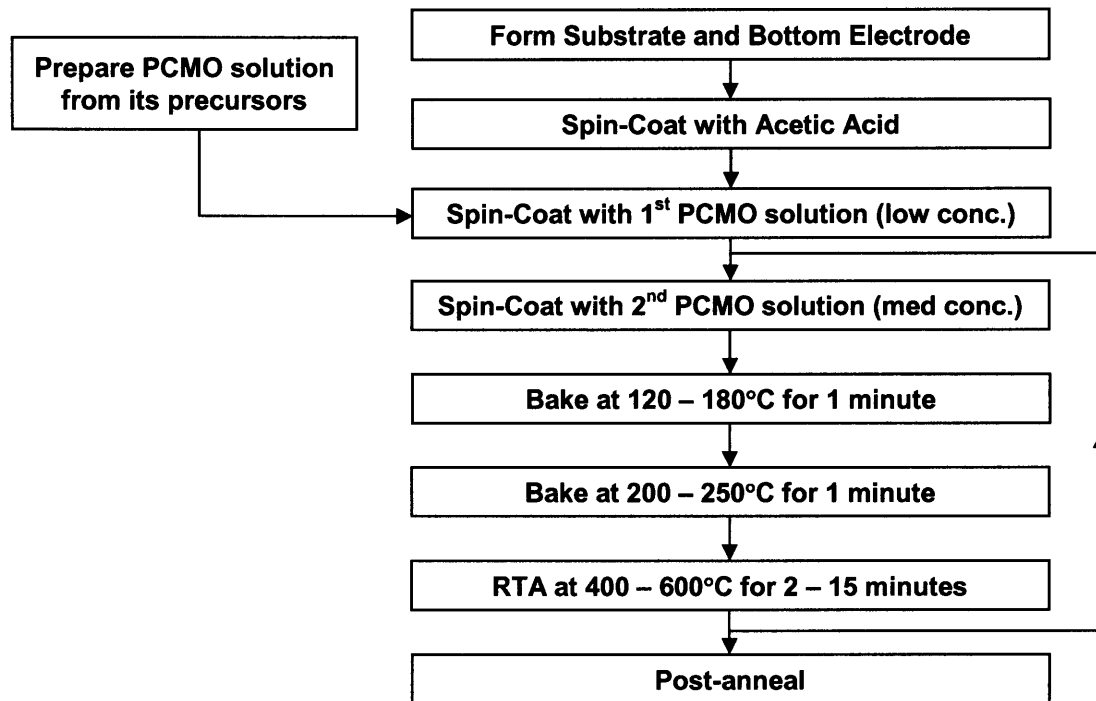
Figure 22 PCMO spin-coat deposition processes as proposed in [62]

After the substrate and the bottom electrode have been prepared, the surface (of the bottom electrode) is spin-coated with acetic acid at 1500 – 3000 rpm for 30 seconds. After preparing PCMO solution from its precursors (i.e. $Pr(CH_3CO_2)_3.H_2O$, $Ca(CH_3CO_2)_2.H_2O$ and $Mn(CH_3CO_2)_3.2H_2O$), a low-concentration (i.e. 0.05M) of the solution is spin-coated at the same speed and time. Then, another 0.25M PCMO solution is injected while the wafer is spun at 500 rpm and followed by the same treatment as earlier (1500-3000 rpm for 30 seconds). Then, the substrate is baked at 180°C and 230°C, each for 1 minute. RTA (rapid temperature annealing) is then performed at 500°C for another 5 minutes. From the injection of 0.25M, the subsequent processes can be repeated for about 2-6 times to achieve the desired thickness. Finally, post-anneal treatment at 500°C to 600°C for 1-6 hours in dry clean air concludes the fabrication. The

PCMO film deposited via spin-coat process can be grown from 40 to 500 nm thick with additional voids having diameter less than 50 Å potentially formed between the bottom electrode and the film if the baking and annealing steps do not proceed uniformly.

After all features that are embedded in dielectric layer have been deposited, the planar surface of the device is further smoothened via CMP (chemical mechanical polishing). As a final step, bit lines perpendicular to the trench and connecting the resistive components (4,5,7) of the same row is deposited using the same method as copper metallization, i.e. electrodeposition.

4.4.2 Fabrication of 1R1D cell
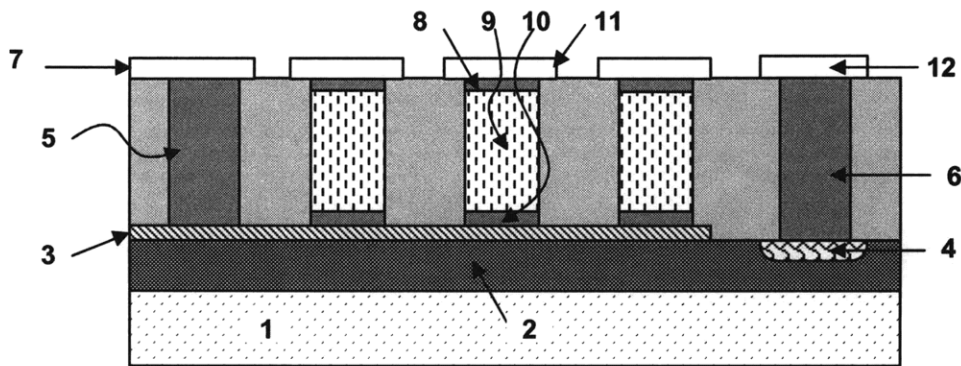
Legend:
1. Substrate
2. *P*-type region (Word Line)
3. Ti (Schottky diode)
4. *P+* Si (Contact point for WL)
5. Cu (Interconnect of Diode and Contact Window 7)
6. Cu (Interconnect of Word Line and Contact Window 12)
7. Cu (conducting line)
8. Pt (top electrode)
9. PCMO (resistive component)
10. Pt (bottom electrode)
11. Cu (Bit Line)
12. Cu (conducting line)

Figure 23 Schematic lateral view of 1R1D cell along with its material components as proposed in [53]

Apart from the obvious difference (i.e. diode replaces transistor), the fabrication techniques of 1R1D structure is nearly similar. In the above illustration, a Schottky diode is used in which metal Ti is in junction with *p*-Si (2). The resistive component (8-10) is arranged in parallel with the Schottky diode. p^+-Si is formed to improve the transport along the WL (2) and interconnect (6) junction. Conducting line (7, 12) perpendicular to trench is deposited on top of the interconnect in order to make WL (2) and diode (3) electrically connected with the external circuits.

4.4.3 Summary of the RRAM materials and fabrication

Table 2 summarizes the fabrication techniques and required materials to produce an RRAM cell. Even though the list is extracted from the discussion on 1R1T cell, the thesis will consider 1R2T cell as its main subject. The idea of focusing only on 1R2T cell is based on consideration that this architecture is more robust towards the programming error (e.g. voltage stress phenomenon), despite compromising the size. This architecture will also be the focus of the subsequent discussion, particularly during the cost analysis (Section 5.4.1).

Table 3 Fabrication Method and Materials for the memory cell[52,53,62]

Function	Materials	Precursors	Process and other details
Substrate	Silicon	Si wafers	Pre-fabricated
Bottom and top electrode	Pt	Bulk Pt	Sputtering and patterning
Dielectric	SiO_2	O_2 gas	Dry oxidation
Gate Transistor (metal dielectric) (metal gate)	HfO_2 TiN	Hf metal, O_2 TiN target	MOCVD Sputtering
Metallization/ interconnect	Cu	Bulk Cu	Electrodeposition
Photoresist	DNQ-Novolac	DNQ-Novolac	Lithography (patterning)
Diffusion barrier	TaN	$Ta[N(CH_3)_2]_5$ and NH_3	CVD
Resistive element	$Pr_{0.7}Ca_{0.3}MnO_3$	$Pr(CH_3CO_2)_3 \cdot H_2O$, $Ca(CH_3CO_2)_2 \cdot H_2O$ and $Mn(CH_3CO_2)_3 \cdot 2H_2O$ in acetic acid solvent	Spin-coating, baking and annealing

4.5 Other application of CER effect: memristor

There are several areas for which colossal electroresistance effect may have potential application; the most notable and interesting one is the device called memristor. Before proceeding further, it is of precautionary that the memristor discussion presented here would not be in depth; only topics that are relevant with CER and RRAM will be discussed.

4.5.1 Background

In 1971, Leon Chua[64] reasoned from symmetry arguments that there should be a fourth fundamental element, which he called a memristor (short for memory resistor). Despite the interesting and valuable circuit properties that this then-hypothetical device might display, until recently when the CER effect and the RRAM technology emerge, no one had presented a physical example of a memristor. Chua noted that from four fundamental circuit variables,

namely electric current i, voltage v, charge q and magnetic flux φ, six (i.e. $_4C_2$) different mathematical relations connecting any two of them can be constructed. Three of them define the classical passive circuit element relations, i.e. resistors, inductors and capacitors, while two of them are derived from the classical electromagnetic theory, i.e. current-charge and voltage-flux relationship. Illustration of these variables is given in Figure 24.
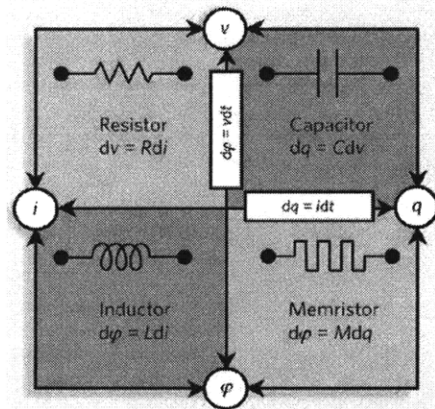
Figure 24 Schematic relationships among four circuit variables, i.e. current, voltage, charge and magnetic flux[65]

So far only five of six relationships have been defined and the expression on the bottom right box completes it. This box displays the 'missing element' Chua had foreseen almost four decades ago: the memristor, defined in simple expression as $d\varphi = M \cdot dq$ where M is the memristance. In the case of linear elements (or constant M), memristance is identical to resistance. Things become more interesting when M is a function of q, resulting in a nonlinear circuit element with hysteretic-like i-v characteristics. It was shown that no combination of nonlinear resistive, capacitive and inductive components can duplicate the circuit properties of a nonlinear memristor while inclusion of active circuit elements such as amplifiers may help.

4.5.2 Memristor versus RRAM

The hysteretic I-V curve of memristor looks identical with one that is observed in RRAM. Indeed, it was discussed in a literature[66] that the resistance switching behavior of RRAM is an example of physical manifestation of memristor; in other words, the 'missing' element has been found. In that case, a curious mind might ask that if the concept of memristor was born many years ago, why it takes more than three decades for such idea to materialize. The main possible answer to that question would be that the unique characteristic of memristor requires coupling of the

several atomic-scale order parameters which defines the internal state of the device with the externally-induced electronic transport. The coupling effect could only become relevant if the active region of the device shrinks to a dimension of few nanometres[66]. Such nanoscale miniaturization requires processing techniques that are only available from the past 10-15 years.

Now that the relationship between memristor and RRAM has been established, other questions might arise: do these names refer to the same thing? If not, what are their differences? The answer to the first question is no. RRAM is seen as a form of memristor's application exploiting its memory characteristic for storing information in a non-volatile way. The memristor itself is a basic circuit element whose resistance-switching characteristic can be applied extensively, not only limited to information storage purpose. In his subsequent papers, Chua noted that the memristive system[67]—defined as any class of two-terminal devices whose response (e.g. resistance) depends on the coupling of the internal state of the system with the external stimuli—can actually be observed in many systems. Examples include the temperature-dependent thermistor[68], the adaptive behavior of unicellular organism such as amoebas[69], etc. It has also been suggested that the concept of memristor can be extended for construction of artificial neuron system due to the fact that a synapse in the brain is strengthened whenever current flows through it, much the same way as current flowing through a memristor lowers resistance.

If RRAM is used as the sole evident of the realization of memristor whose characteristics were described by Chua in his original paper[64], then the claim that the memristor has been 'found' is highly arguable. That is so since by its semantic and technical definition the memristor is more general than a mere data storage device which RRAM is. Nevertheless, the pursuit on making a memristor that obeys Chua's description still continues. Among all the companies attempting for a breakthrough in this field, only HP Labs has publicly announced that the memristor is one of its long-term target; others choose to focus on a much short-term target, i.e. the RRAM cell. A more elaborate discussion on the development of memristor in HP Labs is presented in Section 5.5.1.

4.5.3 Further application: *mem*-circuit elements

The realization of memristor in the form of RRAM has brought interesting light to the scientific world. A particularly interesting expansion of memristor's idea is formulated in the early 2009[70].

It was first reported in 2001 by a group of Korean scientists that a hysteretic capacitance-voltage (C–V) behavior was observed in metal/SiO$_x$/Si structure containing nanocrystals[71]. The same hysteretic C-V behavior was also observed in the MIM structure—with perovskite oxides (PCMO) as the insulator—and subsequently referred to as EPIC[72] (electric-pulse-induced capacitance change) effect by the same group that suggested the name EPIR (electric-pulse-induced resistance change) (Chapter 3). These hysteretic curves are shown in Figure 25.
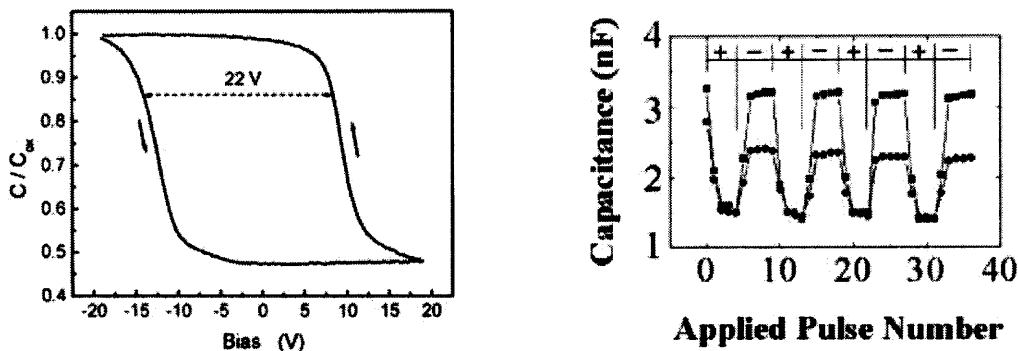
Figure 25 Capacitance-switching observed in (left) nanocrystals-containing metal/SiOx/Si system (voltage sweep at 2V/s)[71] and (right) PCMO (input signal of 18 V within 103 ns)[72].

Using these results, Chua et al [70] then revised and expanded his memristor concept to include these experimental findings. The revised theory proposed that the memory characteristics should not be limited to resistors; other passive circuit elements can have their mem- counterparts, i.e. mem-capacitors and mem-inductors (Figure 26).

Memristor   Memcapacitor   Meminductor
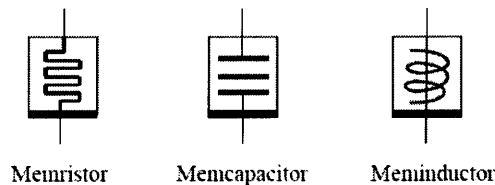
Figure 26 Suggested circuit symbols for the novel mem-devices[70]

The characteristics of these unfortunately named elements have yet to be understood. Yet, Chua et al in the same article[70] envisioned the applicability of these elements in the neuromorphic computer architectures which may further lead to a variety of new applications in the future.

41

Chapter 5
Commercialization potential of RRAM

5.1 Market analysis

The properties of RRAM discussed in Chapter 2 suggest that RRAM has a potential to become the candidate for the holy grail of the memory industry, the universal memory. The incomplete understanding of the physics behind it, the complexities of the fabrication processes as well as the current market competition have however indicated otherwise. In addition to that, the history of several emerging technologies—as well as the matured ones—often taught us that targeting a specific market instead of aiming for a jack-of-all-trades role often proves to be the strategy for a successful market penetration and continuous growth, allowing the technology to develop and survive. Considering its memory characteristics, RRAM is predicted to be more promising and have higher chance to succeed if it targets the non-volatile memory market, or the "storage" industry. This is so since the data retention and the ability to have multilevel states would unfortunately be of no use if it is directed toward the volatile memory. Nonetheless, as is explained earlier, application in the volatile memory is still possible thanks to the rapid switching time characteristics.

5.1.1 Current memory market situation

Before deciding the appropriate commercialization strategy for RRAM, market analysis shall be performed to understand the current market situation and what important factors that must be taken into consideration. Report on global sales for the computer chips industry in Year 2007, as shown in Figure 27, indicates that the total share of memory chips (both volatile and non-volatile) captures approximately 23 – 25% of the total global sales. Within this figure, DRAM (the most dominant volatile memory) takes up about 12% of the total share while Flash (the most dominant non-volatile memory) enjoys 9% of this figure.

GLOBAL SALES FOR SELECTED CHIPS
(Ranked by 2007 estimated sales, in billions of dollars)

E-Estimated. F-Forecast. DRAM-Dynamic random access memory. DSP-Digital signal processors.
SRAM-Static random access memory.
Source: Semiconductor Industry Association.
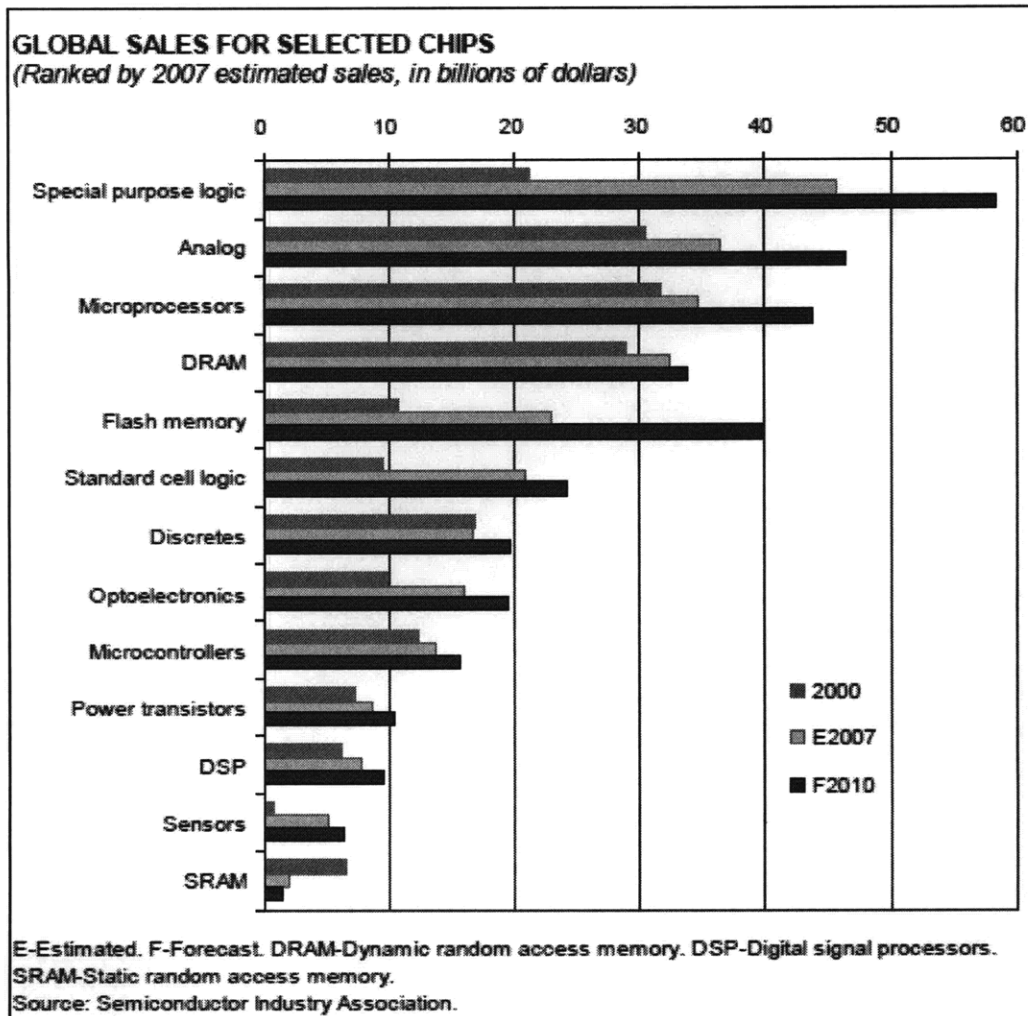
Figure 27 Global Sales of Computer Memory in 2007[73]

The above sales figure, however, provide an incomplete view on how the industry has been performing over the years. The ramification of the 2009 global financial turmoil to the semiconductor industry can be much clearly understood using the market growth data, as illustrated in Figure 28. In this chart, the economic slowdown can be easily identified from the declining growth of DRAM and Flash memory market since the end of Year 2006.
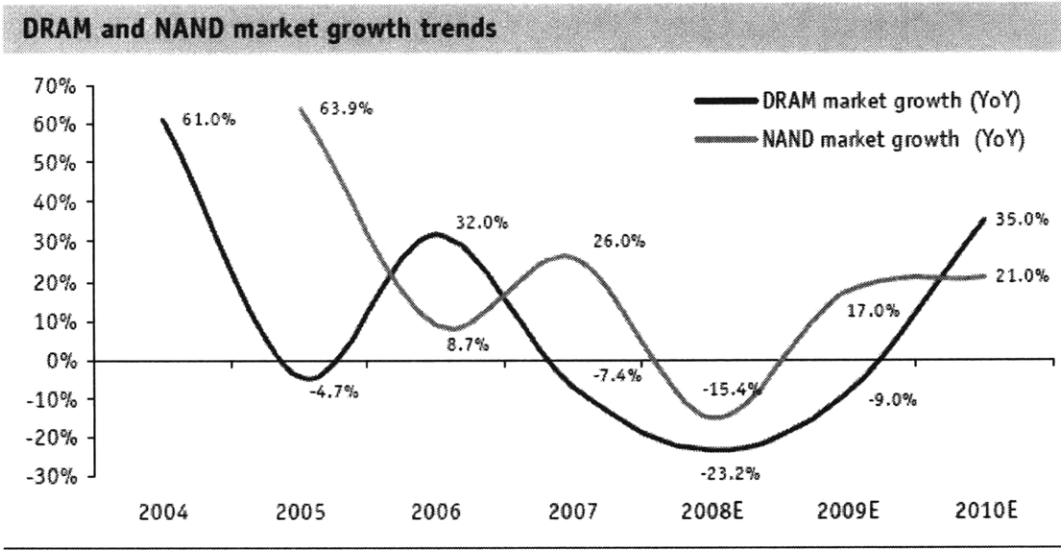
Figure 28 DRAM and NAND Flash market growth (2004-2008)[74]

Not only the market growth of Flash memory shows a plummeting trend but its selling price has also been slowly decreasing, albeit fluctuatively, as depicted in Figure 29. Several dynamic factors contribute to this development, namely the over-supply from the increasingly crowded market and the surge in price due to initially excessive demand. In addition to those, customers have also the tendency to refrain from going in for any further upgrades, on account of sufficient capacity in their existing flash cards and USB drives.
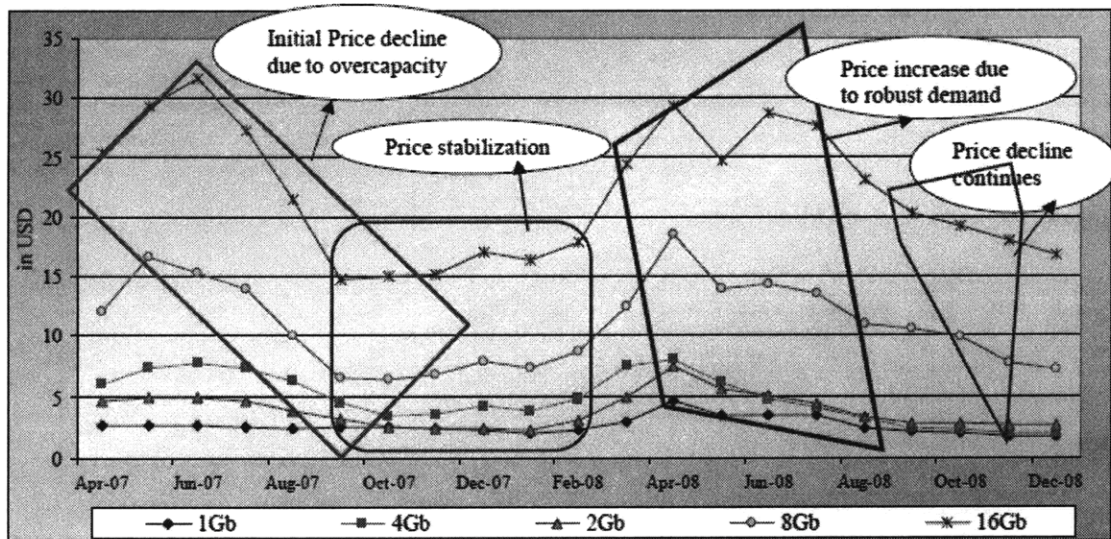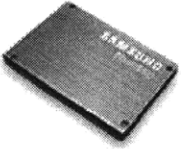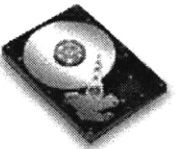
Figure 29 Market price fluctuation of Flash memory[75]

44

Taking these conditions and the fact of the current fierce competition among the Flash memory players into consideration, it would seem to be unlikely for RRAM to position itself as an alternative for the next non-volatile memory technology. One of the plausible logical responses is then to acknowledge a premature death of the non-volatile memory RRAM. However, one certainly should not draw a decision based on an incomplete view of the whole industry. In recent years, another memory-related market segment has shown significant growth, particularly due to the increasing demand of portable computing devices. The segment that it was referring to is the solid-state drive market.

5.1.2 Solid-State Drive and its market

Solid-state drive is a data storage device that employs solid-state memory to store persistent data. An SSD device is considered as technological substitution for the hard-disk drive (HDD) found in a typical personal computer. The use of the term "solid-state" refers to the fact that this device works based on the solid-state electronics, the opposite of HDD in which electromechanical mechanism underlies its core principle. With no moving parts, SSD is seen to be more mechanically robust than HDD and produces less noise (without the cooling fan). In addition to that, the SSD also displays various technological merits over the currently dominant HDD. Using the product's data from Samsung, feature comparison between these two is presented in Table 4.

Table 4 Feature comparison of HDD and SSD (Samsung)[76]

2.5" SATA 3.0Gbps SSD		2.5" SATA 3.0Gbps HDD
Solid NAND flash based	Mechanism type	Magnetic rotating platters
64GB	Density	80GB
73g	Weight	365g
Read: 100MB/s, Write :80MB/s	Performance	Read: 59MB/s, Write: 60MB/s
1W	Active Power consumption	3.86W
20G (10~2000Hz)	Operating Vibration	0.5G (22~350Hz)
1,500G for 0.5ms	Shock resistance	170G for 0.5ms
0°C to 70°C	Operating temperature	5°C to 55°C
None	Acoustic Noise	0.3 dB
MTBF >2M hours	Endurance	MTBF < 0.7M hours

Due to its distinctive features, Solid-State Drive (SSD) has been widely applied in military and industry products. With the beginning of adoption by PC, SSD soon expanded its applications to various consumer electronics, such as notebooks, mobile Internet devices (MID) and smart phones, for which it is predicted that over 90% will have embedded Flash memory by 2012[77]. MacBook Air by Apple is an example of the recent electronics employing SSD as its storage component. Although the lower acceptance in low-price PC and boosting NAND Flash price since 2008 has slowed its growing momentum, SSD is still perceived as a potential product.

In such a growing market, the competition certainly becomes tougher. As recent as June 2009, many vendors aggressively introduced new SSD products during the global-scale exhibition (Computex) in Taipei[77]. Kingston has launched Now-M and Now-E SSD series for commercial PC and servers, both characterized by their effective read/write and I/O performance. San Disk also claims that their new products, pSSD P2 and pSSD S2, can sustain up to 9000vRPM (virtual revolutions per minute) and perform much effectively with the 320MB SRAM that supports the writing function. Other Taiwanese module houses, such as PQI, have also emerged to introduce series of SSD products to the market.

Despite its significant progress, SSD carries a major problem that can become obstacle for its future growth: its selling price. Below is the comparison of the price per GB of both as of 2008. It is clearly shown that the SSD is still an order of magnitude more expensive than the HDD.

Table 5 Price comparison of HDD and SSD (as of 2008)[78]

	DRAM	SSD	HDD
Budgetary Cost	$100/GB	$35/GB	$5/GB
Power Consumption	463 Watts	2.5 Watts	12 Watts
Random I/O	1,000,000 IOPS	7,000-50,000 IOPS	350 IOPS

The SSD sales projection (Figure 30) displays a trend in which increasing growth yet low demand are its market characteristics. Such might be a result of the combination between its technological feature and cost factor. Its excellent technological feature significantly drives its growth rate while the high unit cost maintains its absolute selling figures at relatively low level (about 10 million units in 2009).
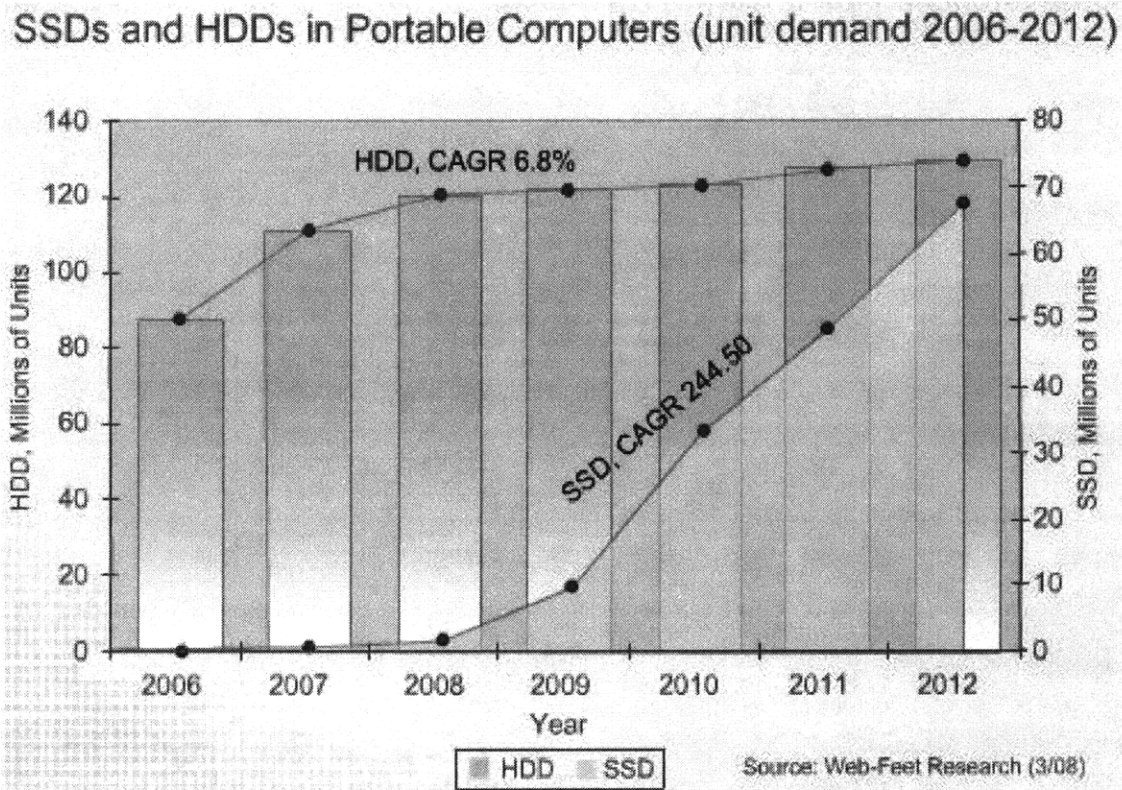
Figure 30 Projected sales figures of HDD and SSD[79]

Acknowledging the relatively stable SSD market condition and the fact that Flash memory virtually serve as the sole-player—a semi-monopoly situation—in the current SSD market, RRAM may consider this as a niche market to go for its future application. Should this strategy be chosen, RRAM will be playing against the following players.

1) The existing HDD vendors whose sales performance is shown in Table 6 below

Table 6 World Wide Hard Disk Drive Revenues, by Vendors (Ranked by 2007 Revenue)[80]

Company	Revenues (US$M)			MARKET SHARE (%)	
	2006	2007	% Chge	2006	2007
Seagate*	10.3	12.3	18.9	34.9	37.9
Maxtor*	1.4	…	…	4.7	…
Western Digital	4.9	6.6	34.5	16.6	20.4
Hitachi	4.9	5.6	14.1	16.5	17.2
Samsung	2.5	2.8	12.6	8.5	8.8
Fujitsu	2.6	2.7	1.9	8.8	8.2
Toshiba	2.8	2.3	(15.6)	9.3	7.2
Others	0.2	0.1	NA	0.7	0.3
Total	29.6	32.4	9.5	100.0	100.0

*Seagate includes Maxtor shares in 2007. NA—Not Available. Note: % change figures based on unrounded data
Source: IDC's May 2008 forecast report

2) The current players in the Flash memory market as shown in Table 7 below. It is worth to note that these vendors are ranked based on their performance in the total Flash memory market. Some of them have joined the SSD race (e.g. Samsung and Intel) while the remaining companies chose to maintain a steady focus on their current technology.

Table 7 World Wide Flash Revenue by Vendors (as of May 2009)[81]

Company	1Q09 Sales (US$M)	1Q09 Market Share (%)	1Q09 Sales QoQ Growth (%)	4Q08 Sales (US$M)	4Q08 Market Share (%)
Samsung	750	36.0	0.7	745	34.9
Toshiba	740	35.4	0.3	738	34.5
Micron	200	9.6	-6.5	214	10.0
Hynix	171	8.2	-22.3	220	10.3
Intel	155	7.4	3.3	150	7.0
Numonyx	70	3.4	0	70	3.3
Total	2,086	100.0	-2.4	2,137	100.0

* 1Q09 Foreign Exchange Average Value: 1 USD = 93.6 JPY, 1 USD = 1414 KRW

** 4Q08 Foreign Exchange Average Value: 1 USD = 96 JPY, 1 USD = 1363 KRW

Source: DRAMeXchange, May 2009

5.2 IP landscape

Before developing any business strategies, it is useful to understand the intellectual property (IP) landscape of the RRAM technology. Since the first research publication of RRAM in early 2000s, more than 40 patents have been filed by and issued to different types of assignees: industrial companies, research institutes, universities and individuals. The nature of claims of these patents varies from one to another, nevertheless, they are to be categorized into the following groups: architecture of the memory cells, fabrication method, testing and characterization and other application (memristor, crossbar latch, etc). The patents are also to be grouped according to their assignees. As one can see from the complete list of patents on Appendix 1, there are only few institutes or companies that have filed and successfully been assigned to these patents. In order to provide a brief yet complete view of these patents, the statistical distribution of RRAM-related IPs according to their category and their assignees are shown in the following pie charts.

Distribution of RRAM Patents (category) 2001-2009

- Memory architecture and design
- Fabrication method
- Testing, characterization
- Others (multilevel memory, crossbar memory)
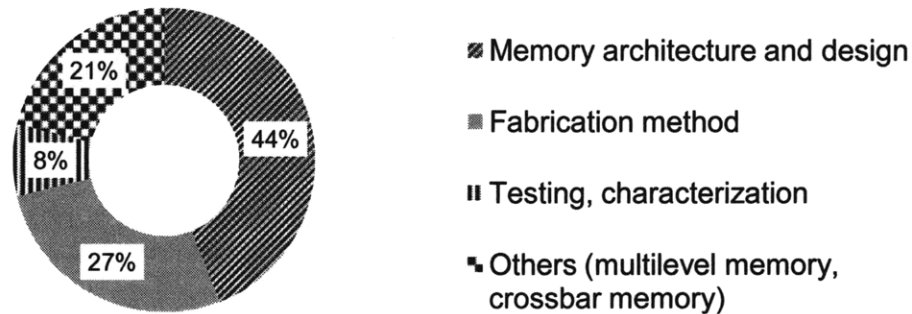
Distribution of RRAM Patents (assignees) 2001-2009

- Unity Semiconductor
- Hewlett Packard
- Sharp (US)
- Sharp (JP)
- Univ. Houston
- Winbond (TW)
- Others
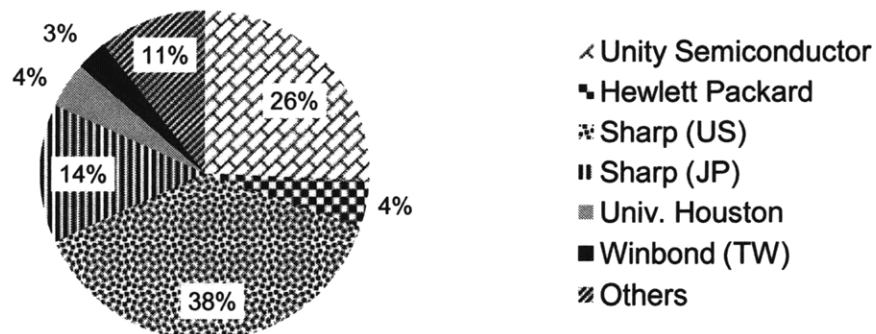
Figure 31 IP statistics based on the field of category (top) and the owners (bottom)

Among these patents, there are several important ones that mark the development of RRAM—from the laboratory testing to its practical implementation—as are listed in Table 8. Some of these patents have been discussed in this chapter, e.g. Patent No. 6,946,702 and 7,016,222.

Table 8 Key Patents and Their Owners

Patent No	Category	Assignee(s)	Title
7,082,052	1	Unity Semiconductor	Multi-resistive state element with reactive metal
7,016,222	1	Sharp Kabushiki, Japan	Non volatile semiconductor memory device
7,009,278	1	Sharp laboratories of America, Inc	3D RRAM
6,586,965	4	Hewlett Packard Development Company	Molecular crossbar latch
6,473,332	1	University of Houston	Electrically variable multi-state resistance computing
7,098,043	2	Sharp Laboratories of America, Inc	PCMO spin-coat deposition
6,841,833	1	Sharp laboratories of America, Inc	1T1R resistive memory
6,856,536	1	Unity Semiconductor	Nonvolatile memory with a single transistor and resistive memory element
6,673,691	3	Sharp Laboratories of America, Inc	Method for resistance switch using short electric pulses
6,204,139	3	University of Houston	Methods for switching the properties of perovskite materials used in thin film resistors
6,664,117	3	Sharp Laboratories of America, Inc	Method for resistance memory metal oxide thin film deposition
6,946,702	1	Winbond Electronics Corp, Taiwan	Resistance random access memory

* Category: 1-- Memory architecture and design; 2--Fabrication method; 3--Testing & characterization; 4--Others development

5.3   Business strategies

There are several business strategies that can be derived from the patents landscape. This section attempts to deduce the strategies that might be taken by the companies that own these patents. It is worth to note that up to the time this thesis is written, almost none of the patents' owners have started the production, let alone the distribution, of RRAM or CER-related memory. Therefore, the three distinct strategies presented here only serve as a mere forecast of what might happen.

1)   Continuous research and development to improve manufacturability and performances of RRAM such that it can perform as good as, if not better than, Flash memory. Universities and research institutes have partaken into this strategy. Their business strategy comprises mainly the IP business model in which filing the patent as early as possible may result in certain amount of profit via patent-licensing to other company. With a sufficient capital, they may also build a start-up company based on their patent.

2)   Aggressive approach to participate in the solid-state drive race. The analysis given earlier on the SSD market indicates that it would be a good place to start for RRAM. Considering the level of maturity of Flash memory, RRAM company should take a rather aggressive approach on the development and mass-production of a working RRAM-based SSD. Tremendous amount of investment for the research and the manufacturing is required to ensure that the RRAM technology be mature only within a short period of time (about 5 to 10 years). Two kinds of business models can be derived from this strategy: Fab-less or Fab-lite business model (Figure 32)
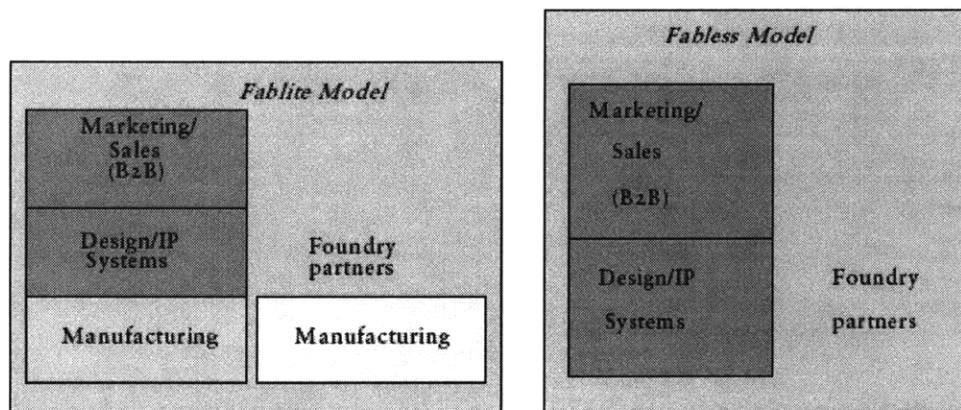
Figure 32 Fab-lite and Fab-less Business Model[82]

51

A fab-less company is one that designs and markets its own devices yet has no foundry plant or wafer production capability. In practice, a company is considered as fab-less if more than 75% of its manufacturing requirement is out-sourced to a third-party foundry. On the other hand, a fab-lite company still performs an operating manufacturing activity but has a policy of outsourcing part of its manufacturing requirement to merchant pure-play foundries.

Looking at the number of patents and their varieties, the Sharp Labs (both U.S. and Japan) and the Unity Semiconductor would most likely adopt this type of strategy. Fab-lite business model is considered more practical for this solution as Sharp Labs and Unity have already owned many different patents on the fabrication processing of such completely new device, RRAM cell. Start-up companies without any RRAM's IPs might also join the race and run their business using one of these strategies. However, they will incur higher production cost from the IP-licensing since most of the patents have yet to expire, i.e. they were issued only recently (less than 5 years).

3)   Creation of its own 'blue ocean'. Instead of aiming for a seemingly saturated memory market, a company can choose to participate in continuous novel research while in the same time developing a working prototype of the device that no other companies are pursuing. In the marketing world, this strategy is also known as the blue ocean strategy. By creating a completely new market and consequently a technological entrance barrier for other competitors, the company can enjoy a virtually undisturbed market share and has the potential of long sustainability.

HP Labs seems to go with this direction. Their activities on understanding and developing the application of RRAM concept in novel device, particularly the memristor are distinctive and only pursued currently by this company alone. This strategy, however, will require a more tremendous amount of capital to succeed. Furthermore, as the nature of these devices is not clearly understood for now, there are various degrees of risks that may be involved.

5.4   Cost modeling

After discussing the possible strategies taken by any of the companies owning one or several RRAM patents, the thesis will attempt to develop a simple cost model based only on one of these strategies. Among the three strategies above, the second strategy that incorporates certain level of manufacturing aiming to penetrate the SSD market has a relatively moderate level of complexity for the cost modeling and will thus be analyzed in this section. Furthermore, a fab-lite semiconductor company owning significant number of IPs is assumed; i.e. no licensing cost will be considered.

Other assumption includes one that considers the company to perform the device fabrication from the designing-stage to the materials processing stage, but not including the testing, assembly and the packaging. This last stage before the finished final product will be outsourced to other company; the same way one will outsource the semiconductor substrate from the foundry supplier. Nevertheless, the cost incurred from outsourcing will be included as a certain percentage of the variable cost. And to begin the modeling, the schematic flow of the supply chain is illustrated in Figure 33.
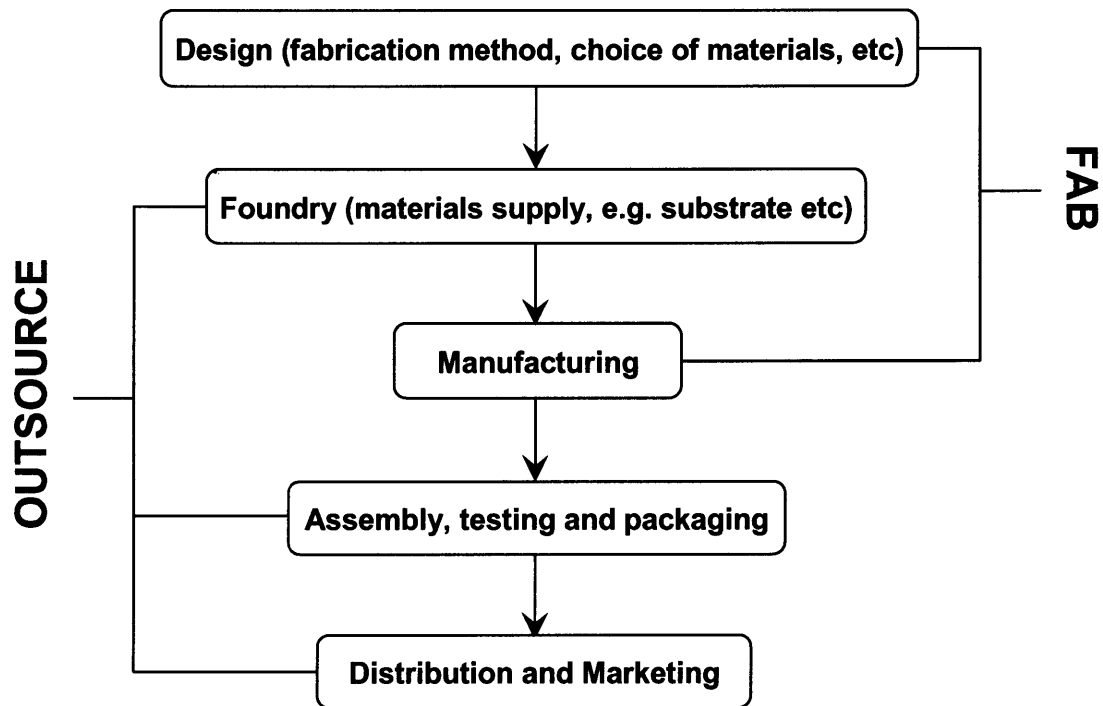
Figure 33 Production chain as derived from Strategy 2 (Section 5.3)

53

5.4.1 Preliminary assumptions

The followings are the preliminary assumptions employed in formulating the cost model. They consist of the assumption about the product specification, the processes, the production target and the production line.

1) **Final Product**

   A 64-GB SSD memory cell is assumed. The cell would incorporate multilevel switching with 4 bit/cell and using the 1R2T (Section 4.4.3) architecture covering an area of 16×16 mm^2. The 45-nm half pitch technology node is employed and the device is fabricated on top of a 12-inch Si wafer. These numbers are estimated from the various current SSD memory cells, e.g. Intel and HP[83].

2) **Manufacturing process and production line**

   The manufacturing processes include those which have been discussed in Section 4.4. For the feature dimension, a 45-nm technology node is assumed. In order to achieve this scaling without going for the EUV or electron beam lithography, double-patterning water-immersed lithography (NA = 1.35) is assumed. The drawback, however, is that the double-patterning technique requires almost twice as many masks as the conventional lithography.

3) **Production target**

   Using the sales figure from the Flash-SSD market in 2009 (Figure 30), the model will assume a target of 2% to 5% market share corresponding to about 600,000 units produced annually. Production plan is executed in 5-years period, beginning with the first year in which near-full capacity is assumed and in this case production volume is taken to be 500,000 units. Such assumption is applied to allow for flexibility in making any necessary adjustments as well as to gain the efficiency through the learning curve.

5.4.2 Description and outline of the cost model

Given a wafer of diameter d (mm) and target IC size S (mm^2), the gross number of die can be estimated using the die-per-wafer expression[84]:

$$DPW = d\pi\left(\frac{d}{4S} - \frac{1}{\sqrt{2S}}\right) \qquad \text{... Eq (3)}$$

Substituting the 12-inch wafer and the IC area of ~250 mm^2, the number of dies (or the number of one unit of the 64-GB cell) is 241 dies/wafer. Before slicing them into dies, these wafers are grouped in a 'lot' containing 40 of them which then will undergo series of fabrication steps (Section 4.4) within one production line. Based on the available production time, the scrap rate from each fabrication step, and the production capacity—all of which are listed in the spreadsheet given in Appendix 2—there are four production lines working in parallel. One technician is assigned to operate similar equipment from two production lines; while a supervisor will handle only one production line.

The accounting lifetime of each of the equipments is estimated as 10 years and their nominal value would be discounted at 15% per annum. The price of each equipment is estimated from the average value of the price obtained from different manufacturers and semiconductor's industry database (e.g. Applied Materials, Sematech, etc)[85]. The rate of materials' consumption is obtained from simple geometrical structure of the memory cell which then is scaled to get the consumption rate per wafer. The unit price of these materials is obtained mainly from the catalogue of two semiconductor chemical suppliers, e.g. Sigma-Aldrich and Strem Chemicals[86]. Other input variables including energy consumption per each process, planned downtime, the infrastructures (building, clean rooms etc), and several others are extracted from various database and references.

Using the spreadsheet tool designed by the MIT Materials System Laboratory, these inputs are processed to result in the summary of the production cost. In addition to the Cost Model spreadsheet, the MIT Materials System Laboratory developed another spreadsheet tool, the Decision Analysis spreadsheet, which may assist in providing a more profound understanding of the business strategy and its possible outcomes. The Decision Analysis spreadsheet provides the production plan for various market conditions using the market-related variables such as the growth probability, customer's utility function, market uncertainties etc. As the nature of these variables is volatile and market-dependent, further analysis using this spreadsheet would not be attempted. Nevertheless, a schematic outline describing both spreadsheets, the variables that are involved, and the relationship between the two is illustrated in Figure 34. The part inside the dashed box represents the cost model attempted in this thesis.

55
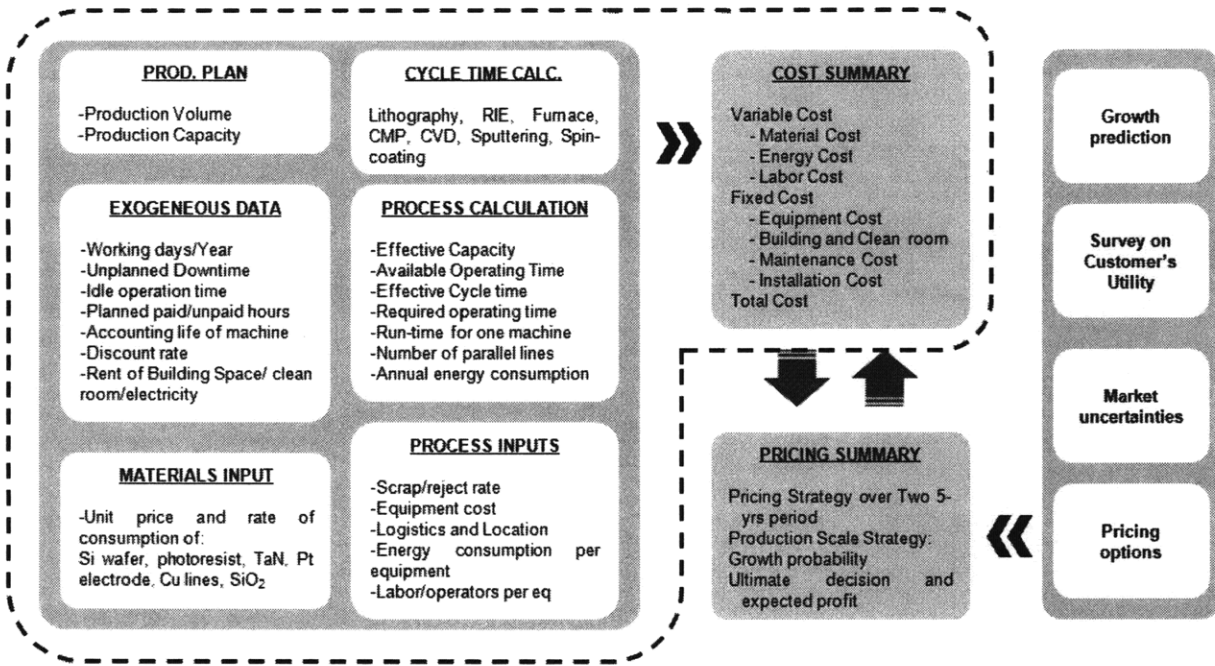
Figure 34 Schematic of Cost Model and Decision Analysis developed via the spreadsheet tools from the MIT Materials System Laboratory

5.4.3 Summary of the cost model

Table 9 below summarizes the cost of production as calculated using the above assumptions.

Table 9 Summary of the production cost based on the assumptions described in Section 5.4.1

Fixed Cost	Variable Cost
$52.38	$81.79
Total Cost	$134.17

Total Cost Breakdown

$52.38; 39%

$81.79; 61%

Ⅲ Fixed Cost
✪ Variable Cost
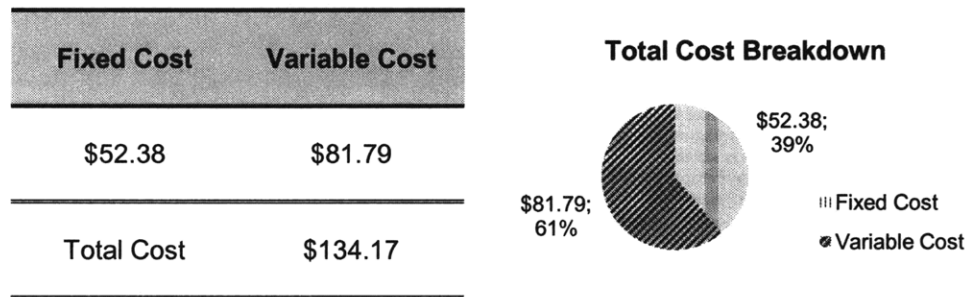
The breakdown of this production cost is listed in the following table.

Table 10 Detailed cost component according to the model derived from Strategy 2

COST COMPONENTS	VALUE	PCTG (%)	PIE CHART
Fixed Cost			
Lithography Machine	$28.69	21.39%	**Fixed Cost Components**
Lithography Mask	$13.40	9.99%	
Other equipments	$5.07	3.78%	
Building and Cleanroom	$2.32	1.73%	
Maintenance	$0.20	0.15%	
Installation	$2.71	2.02%	
Variable Cost			**Variable Cost Component**
Silicon wafer	$59.85	44.61%	
Other materials	$11.41	8.50%	
Energy	$0.02	0.02%	
Labor	$81.79	7.83%	
TOTAL COST	**$134.17**	**100%**	

5.4.4 Cost analysis

The above cost calculation shows that the production cost of a 64-GB cell is achieved at $134.17 with 39% of the total cost is allocated to the fixed cost. Comparing this value with the figures from the latest Intel SSD price on Table 11 (assuming linear price/GB) brings to a preliminary conclusion that the production cost for RRAM cell ($134.17) is significantly cheaper than the SSD price in the market ($250). Such conclusion is nevertheless incomplete since during the calculation of the cost, it is assumed that the company only covers the production chain up to the device manufacturing step.

Table 11 Intel SSD Price List (as of May 2009)[87]

Model	December 2008	February 2009	April 2009
X25-M 80GB	$595	$390	$320
X25-M 160GB	$945	$765	$630
X18-M 80GB	--	--	$340
X25-E 32GB	$575	$415	$390
X25-E 64GB	--	$795	$795

Subsequent production chains that are not considered in the calculation will add significant increase to the production cost. These steps and their respective contribution to the final cost are, but not limited to, those listed in Table 12. It is reminded again that the numbers specified here are collected from various sources and references taken with over-estimated assumption.

Table 12 Expected additional cost components[88]

Subsequent steps	Cost margin (% the manufacturing cost calculated earlier)
Testing	12 – 15%
Assembly and packaging	8 – 10%
Distribution	3 – 7%
Marketing	5 – 10%
R&D (for further development)	7 – 10%
Total	35 - 52%

Finally, the customer's retail price would include the final product cost as well as a certain percentage of profit margins. Considering the fact that the SSD technology is a notably recent one—thus, only few players are currently in this market—the percentage of profit margin is expected to be fairly high; it can be as high as 30% - 45% of the final product cost. This number will be adjusted as more newcomers penetrate into the market, slowly diminishing the oligopoly. Substituting all of these additional cost factors, the retail price of the SSD market is expected to be almost twice as much as the figure calculated from the cost model or about $260. This figure would again is extremely volatile to the market condition, similar to the Intel's SSD significant price drop that occurred only within two months. That being said, the cost-estimation and cost-building of the RRAM-based SSD have been attempted and are done by including many realistic—sometimes even overly pessimistic—assumptions.

5.4.5　Economy of scale in cost modeling

As is similar to the production plan of many other commercial products, this cost model also exhibits an advantageous feature named the economies of scale. This characteristic tells us that the unit cost will further decrease as the manufacturing plant operation is closer to its full capacity. Similar trend is also observed in the case where the production capacity of the manufacturing plant is increased. In other words, the total cost versus production volume curve at a given production capacity will have a hyperbolic trend.

The discussion begins by examining the effect of economies of scale on the cost component under a given production capacity, in this case 600,000 unit annually. Using sensitivity analysis, the unit cost at various production volumes can be calculated, as shown in Figure 35. It can be concluded that the unit cost reduction is contributed only by that same trend found in fixed cost. This shows the agreement between formal definition of cost component—both fixed and variable cost—and the cost-modeling itself; i.e. a constant cost margin called unit variable cost will add to the total cost as the production level is raised while the unit fixed cost—cost that is incurred regardless of the production level—will diminish at the same time.
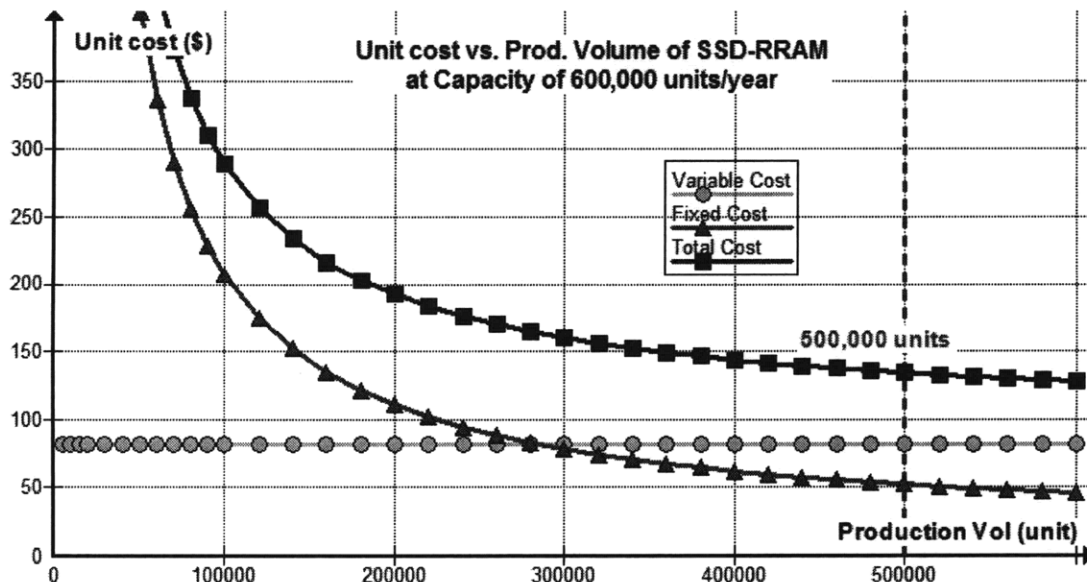
Figure 35 Unit Cost vs Production Volume at the capacity 600,000 units/year

Another aspect of economies of scale is that further cost reduction can be achieved by increasing the production capacity, assuming similar level of operation. In Figure 36, it can be seen that when the plant operates at full-capacity (production volume equals production capacity), the unit cost at this full-operation diminishes as the capacity increases (the decreasing ordinate of each end-point of the curve).
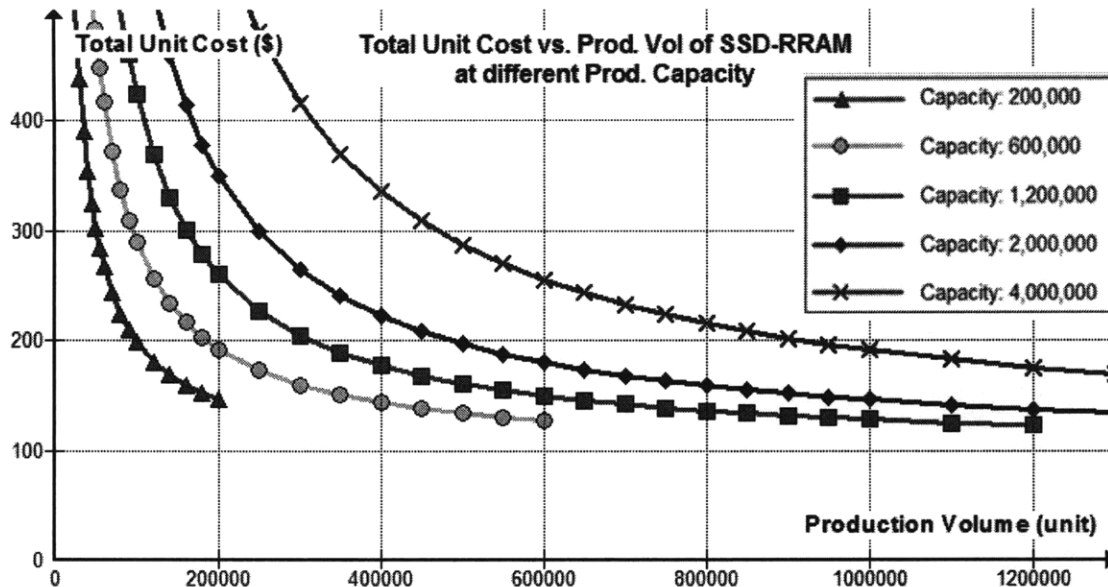
Figure 36 Unit Cost vs Production Volume at different capacities

Such a competitive unit cost certainly comes with a price—a terribly exorbitant one. The lowest unit cost is achieved from the highest production capacity, which in turn requires a tremendous capital investment. In the Cost Model spreadsheet it is shown that to achieve production capacity of 600,000 units, initial investment of over $80 million is a prerequisite. This number could go as high as over $400 million if the capacity is increased to 2 million units, which would certainly involve extremely high risk, especially considering the nature of RRAM technology that is still on the embryonic stage. Other risks that may be involved are related to the emerging SSD market and the potential integration issue with the current technology in the portable computers.

5.4.6 Alternative strategy for unit cost reduction

There are other alternative strategies that may help in reducing the unit cost and thus giving an initial competitive advantage to the company for its market penetration. One of them is to

employ the electron-beam lithography technique, replacing the conventional one. Currently, this technique is considered unsuitable for mass-production because of its low throughput. However, a consortium lead by the Japanese semiconductor industry, D2S (Direct-to-Silicon), announced in 2005 the development of a commercially viable EBL technique that incorporates cell-projection lithography to increase the throughput as described in Figure 37. This technology, however, is reported to be commercially feasible for annual production of below 100,000 units. More details on this Electron-beam Direct Write (EbDW) technology are available in Ref [89].
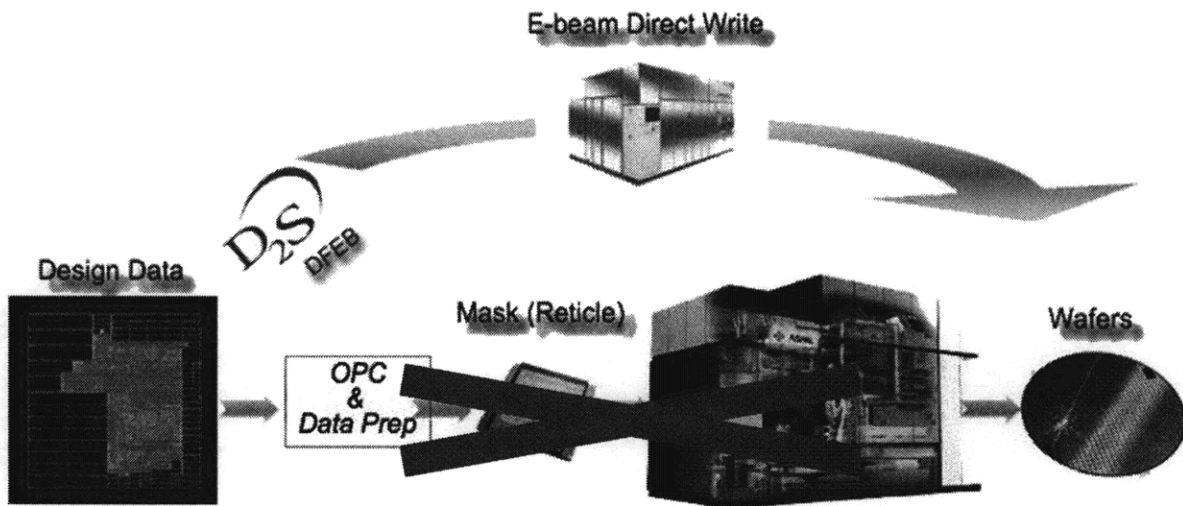
Figure 37 The CP (character projection)-EBL technology developed by D2S Inc [90]

Another plausible cost-reduction strategy is to adapt the fab-less, instead of fab-lite, business model such that the cost of ownership (CoO) of the semiconductor equipments can be eliminated. This strategy, however, removes the advantage of having the economy of scale since the total unit cost would now no longer have a hyperbolic trend (dashed line). Instead, a monotonically increasing trend of unit cost (solid line) is observed as schematically shown in Figure 38.
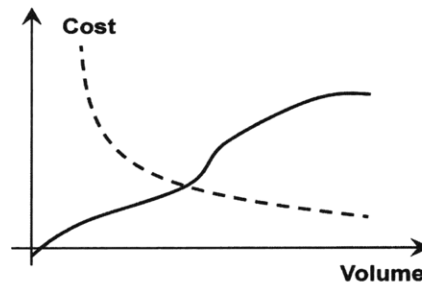
Figure 38 Illustration of cost function (dashed line: with economy of scale, solid line: no economy of scale)[88]

5.5 Recent developments in RRAM industry

Along the preparation and the writing of this thesis, several major developments in the RRAM-related industry have been made and announced to the public. This section summarizes the developments which are mostly originated from the major patent owners discussed in Section 5.2

5.5.1 Hewlett Packard (HP) Labs

As stated earlier, HP Labs publicly stated that they have attempted to develop memristors and sought the possibility of its integration in other application. Presented is the summary of the developments made by this institution. Because a memristor's resistance changes as current flows through it, HP has made the memristor the key element in its overall program to build an entirely new kind of memory using crossbar switches. By using nanoimprint lithography, HP had already perfected a method of fabricating crossing arrays of ultradense perpendicular metal lines—the crossbars. By picking one line from the top array and one line from the bottom, any bit can be directly addressed at the point where the two lines cross, enabling crossbar arrays that today pack 100 Gbits/cm^2. By contrast, Flash was at that time (September 2008) available at 32 Gbits/cm$^{2[91]}$.

For a while, HP was unable to find a reliable material to sandwich between the crossbars, despite several years of experimentation with all sorts of nonvolatile-memory materials. One year before, HP disclosed it was experimenting with organic molecules as the memory element for its crossbars, but the organic material's sensitivity to high temperatures made HP start looking in the inorganic world for a sturdier nonvolatile option. With memristors composed of inorganic titanium oxide, HP believed it finally has the right mix to leapfrog flash and other alternative memory technologies, such as phase-change RAM (PCRAM), with its RRAM.
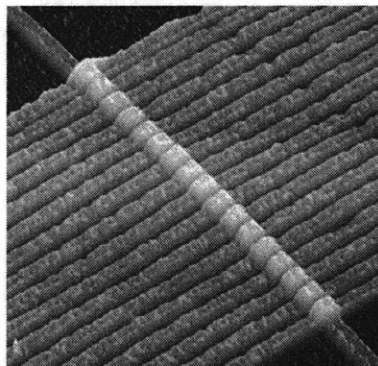
Figure 39 An array of 17 purpose-built oxygen-depleted titanium dioxide memristors built at HP Labs, imaged by an atomic force microscope. The wires are about 50 nm, or 150 atoms wide[92]

HP Labs stated that their prototype RRAM will likely just demonstrate that the memory array itself can be read and written to with external circuitry. The next milestone will be fabricating a crossbar memory array on a conventional silicon chip, with the read/write and addressing circuitry in silicon and the memory array embedded into the metallization layers on top.

5.5.2 Unity Semiconductor

After about seven years since it was founded, Unity Semiconductor issued a press release[93] on May 2009, announcing the launch of its breakthrough memory technology called CMO_x^{TM}. The CMO_x^{TM} technology is said to be based on 'the use of new materials called conductive metal oxides' into the semiconductor process that allows for ionic motion. It is expected to offer storage-class products with 4× the density and 5–10× the write speed of NAND Flash technology. Such feature is envisioned based on the rationale that the CMO_x^{TM} multi-layer cross-point array utilizes a resistance change element—one that lends the memory architecture its passive characteristics. This passive nature would in turn contribute to the high-density memory device and enable the physical stacking of multiple layers of memory. The current design of CMO_x^{TM} uses four physical layers of multi-level cell (MLC) memory and thus is claimed by Unity as 'the world's first passive rewritable cross-point memory array that requires no transistor'[93].
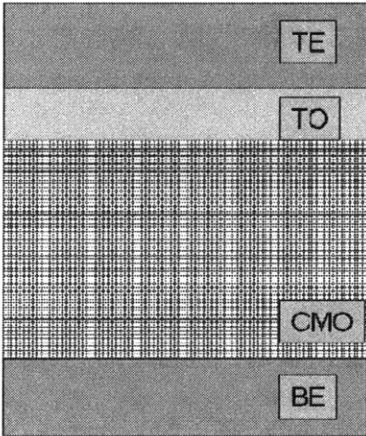
Figure 40 Schematic CMO_x^{TM} cell technology by Unity Semiconductor[94]

To accompany the development of this innovation, Unity stated that they have prepared several business concepts that allow them to manufacture their product at less risk and faster time to

market without involving huge capital investment. The key among these concepts is, as they claimed, the separation of the processing of the front-end-of-line (FEOL) CMOS base wafer from the back-end-of-line (BEOL) memory layer processing. Their BEOL memory concept enables a CMOS logic foundry to be in the memory business without taking significant risks associated with being in the memory business, while their CMOS FEOL strategy allows it to be a moderate follower in CMOS transistor technology, further allowing it to reduce risk and time to market. Using these strategies, Unity has secured partnership with several investors including IM Flash, August Capital, Lightspeed Venture Partners and Morgenthaler Venture[95,96].

This recent development in CER-related non-volatile memory certainly proves that the practical application of such effect is not merely a wishful thinking. Nevertheless, as is with other seemingly utopian stories, this news should always be examined with critical assessment. In their press release, Unity stated that the CMO_x technology that they develop is not an RRAM cell such as is being developed by a few other companies. Looking at the schematic structure of CMO_x, however, might suggest otherwise. Similar to the prototype RRAM device discussed earlier, the CMO_x consists of a CER material (or conductive metal oxide) sandwiched by two metal electrodes. The only different is that there is a very thin oxide layer, called tunneling oxide (TO), between the top electrode and the CER material. While the existence of this layer appears to be the difference of CMO_x and RRAM, the experimental result might nullify this idea. Using high-resolution imaging tools (e.g. HRTEM), few papers [19,97] reported that there exists a thin layer of metal oxide formed under the top electrode upon electroforming (i.e. applying a sufficiently high field to calibrate the resistance state). Even though such layer is only a few angstroms, it may still be an evident to show that the CMO_x technology is nonetheless another name of RRAM.

Another important factor that could prevent the development of CMO_x technology is the legal matters. As reported by Blaise Mouttet[98], one of the people that are indirectly involved in the conception of *mem*-device (Section 4.5.6), there might be some potential conflicts within the patents granted to Unity Semiconductor that now become the basis of CMO_x development. In his self-published article, Mouttet noted that some of the Unity Semiconductor's important patents, e.g. Patent No. 6,870,755, seemed to overlook or conflict with the prior art from other companies such as Hewlett Packard's Patent No. 6,128,214 or Sharp's Patent No. 6,746,910. Such issue did not get much attention previously because none of the major companies targeting the RRAM had

a significant development in the commercialization of this technology. Now that one of these companies managed to secure huge capital investments and proceed into production next year, this legal issue would likely be evoked.

5.5.3 Other companies

4DS, Inc. [99]

An Australian start-up, 4DS, Inc., claimed to have made a major breakthrough in RRAM technology on February 2009, before Unity did the same. The company is now looking for a manufacturing partner to bring its so-called "4DS memory" into mass production. However, unlike Unity, 4DS did not disclose its technological basis. Their CEO, Kurt Pfluger, only stated that, "the 4DS method uses existing semiconductor processes and requires fewer changes to the semiconductor manufacturing equipment, enabling simple manufacturing through a technology that can be scaled significantly farther than NAND or NOR flash". He continued to add that, "4DS' RRAM is a high-capacity, non-volatile memory with fast switching speeds measured below 5-ns, and with an endurance of 1 billion write/read cycles". To prepare for the commercialization stage, 4DS had secured on January 2009 a new round of funding from Poly Plant Project Inc. (Burbank, Calif.), a designer and builder of manufacturing plants for the production of polysilicon.

IMEC (a Belgium-based European research institute) [100]

IMEC's research activities on RRAM mainly focus on investigating the switching behavior of the RRAM cell concept that uses metal oxides as a switching element, and on demonstrating its scaling capability down to 25-nm. The leading candidate is nickel-oxide which has been the subject of study by Samsung at geometries greater than 100-nm. IMEC's study concentrates on three main topics, being RRAM stack optimization (including the choice of top and bottom electrode and of the metal oxide) RRAM cell scaling and RRAM integration in a crossbar RRAM array. Five of the leading memory makers—Samsung Electronics Co. Ltd., Hynix Semiconductor inc., Qimonda AG, Elpida Inc. and Micron Technology Inc.—are involved in the IMEC core CMOS research program and are set to share the cost and benefit from the results of the research.

Chapter 6
Summary and Future Outlook

6.1   Summary

Since the emergence of its experimental discovery in the early 2000, colossal electroresistance (CER) effect has enjoyed a significant level of development in terms of its materials, the theory, and more importantly, the practical application. CER effect, whose variation of names are listed in Chapter 3, refers to the resistance switching behavior found in metal-insulator-metal (MIM) structure during the voltage application of different magnitude and polarity. The insulator material may be perovskite transition metal oxides, binary metal oxides or even soft material. Within the many MIM systems exhibiting this effect, there are some common factors that can be withdrawn from them, namely[101]:

1.   The doping level and the type of carriers highly determine the properties of MIM cell.

2.   The nanoscale phase separation in the system is stimulated by various factors, namely, the level of doping, internal stresses, electric field, magnetic field and etc.

3.   Mobile oxygen vacancies contribute mainly to the charge transport under an applied electric field.

4.   They exhibit a hysteretic I-V curve that corresponds to high- and low-resistance state.

Several theories have been proposed to explain the physical origin of this effect; some of which have been briefly discussed in Chapter 3. Virtually all of these models succeeded only in explaining partial portion of the unique characteristics of CER effect discussed in Chapter 2; most of them fail to explain the multilevel resistance states (Section 2.4) and the unipolar switching (Section 2.1). Such unfinished pursuit has proven to be of trivial impediment for the development of the future candidate of memory technology, the RRAM.

The need to develop a novel kind of memory, particularly the non-volatile one, is based on the notion that the current leading technology, Flash memory, was expected to be obsolete in the near future as the miniaturization trend recommended by ITRS goes beyond 20 nm. The technological limitations encountered by Flash memory at this scaling have been discussed in Section 1.1.3. In addition to the technologically-driven motivation, the RRAM development is

also, even mainly, encouraged by the profit-driven incentives. Section 5.1 illustrates the size of the memory market. Despite the sales decline in most of the computer-related industry due to the 2009 financial turmoil, some of the memory technologies have managed to sustain or even increase their growth; one of them is the solid-state drive (SSD, Section 5.1.2) which would be a feasible target for RRAM to penetrate during its initial commercialization.

Considering the possible financial benefits from the RRAM commercialization, several companies and research institutes have declared themselves joining the race for developing the end-user version of this technology. Many patents related to the RRAM—its design, fabrication methods, integration with more complex devices—have been filed and granted to these companies; some of which are listed in Appendix 1, briefly discussed in Chapter 4 and 5.

Based on the nature and distribution of these patents, three business strategies were derived and elaborated in Section 5.3. From then, a simple model attempting to foresee the production cost of RRAM was built using several assumptions discussed in Section 5.4. It was found that the market price for SSD RRAM, including a reasonable assumption for profit margin, is almost similar to the price of the current SSD Flash. This figure nevertheless can still decline as more players begin to crowd this market segment, driving the price to be at least as competitive as HDD technology. Price decline can also be motivated by the much improved fabrication methods, including the more economic way of producing feature below 45-nm node.

6.2   Current and future outlook

As the research institutes and the industries continue to show progress in the research and development of RRAM, it is expected that the end-user version of RRAM can be introduced to the public within 1-2 years from now (2009). This final section of the thesis will discuss the future outlook from both sides: the technology and the market outlook.

6.2.1   Current state of RRAM development

Discussed earlier in Section 1.2.1, RRAM is projected to be one of the leading candidates for the future non-volatile memory due to its working principle that is based on resistance change. Such is necessary especially in the face of increasingly fierce miniaturization. Several literatures have

reported the experimental verification that supports the RRAM's ability for extreme scaling. The recent data via AFM imaging suggests that RRAM is scalable down to 30 nm[102]. Such scale can be further reduced considering the fact that its key concept lies on the motion of oxygen vacancies which was shown to occur in regions as small as 2 nm[103]. They however do not in any way guarantee the success of RRAM in abiding the ITRS trend since RRAM itself depends on other CMOS components such as the transistor and metallization—both of which are prone to scaling-related reliability problems, such as electromigration and channel's hot carriers effect[104].

In addition to displaying robust feature towards the scaling, the future non-volatile memory technology is also expected to exhibit high storage density (expressed typically in $Gbit/in^2$). The current storage density of NVMs (including HDD and several other memories) is within the range of 80 to 1000 $Gbit/in^2$. This density interval excludes the RRAM technology since, for now, there has not been any working commercialized version of this memory available in public. Some major company, such as HP, has announced that they have succeeded in making a prototype of RRAM with density of 100 Gb/cm^2 (or about 650 Gb/in^2)[91]. Certainly, such claim cannot be validated until their products are out in the market.

The high-density data storage in RRAM is enabled by its multilevel switching feature. However, this may create another issue to the device in which the accompanying software becomes increasingly complex to compensate for a larger BER (Bit Error Ratio). The higher BER requires the writing/reading software to employ an algorithm that can correct errors higher than three bits and detect the condition of more than three bad bits; the most commonly used algorithm is Bose-Chaudhuri-Hocquenghem (BCH) code[105]. This adjustment will certainly affect the cost of the final product but the cost discrepancy is expected to be almost trivial.

The future of RRAM development does not always concern with its properties such as scalability and storage density; in fact, several patents are dedicated to discuss about the materials and its fabrication method[61,62]. Materials selection also plays significant role in the competition among the industry. The three major companies focusing on RRAM—HP, Unity and Sharp—developed their technology using different types of materials. As one of the pioneers in the race, Sharp Labs has the advantage of developing their technology using one of the earliest known CER materials,

i.e. PCMO and other perovskite complex oxides[50]. HP Labs[92] and Unity Semiconductor[93,95] both seem to focus their activities on binary oxides which are arguably considered easier to fabricate than perovskite materials: HP with TiO_2 while Unity with ZnO. Besides perovskite complexes and binary oxides, there is other class of materials displaying the CER effect, i.e. organic materials[13,14]. Nevertheless, their application is still within the laboratory realm as none of the major companies seem to take interest in developing this class of materials.

6.2.2 Future commercialization and technology outlook

During the initial marketing phase of RRAM, the aforementioned companies are likely to target the external portable storage media market since it has the minimum compatibility requirements with the computer unit which it is attached to. As the market grows and the R&D enables more diverse application of RRAM, it is forecasted that RRAM will first attempt to penetrate the solid-state drive (SSD), a segment of non-volatile memory market increasingly more integrated into the portable computers and mobile internet devices (e.g. MacBook Air, iPhone, etc). In addition to that, if Hewlett-Packard (HP) has succeeded in finalizing their memristors architecture, the public might also be able to witness the launch of its first application as early as next year.

While these immediate premises are mainly directed towards the replacement or improvement of the current technology, the case might be different for the long-term goals. Application in novel areas which might even currently be unavailable or not recognized can be derived from the RRAM technology; a possible example of which is the construction of an artificial memristive system from memristors-like components resembling the work of neuron systems as envisioned by Chua et al[69] (Section 4.5.2).

This example would of course bring another possible exotic RRAM-related application, that is, the development of mem-devices (Section 4.5.3)—memristors, memcapacitors and meminductors—whose concepts were conceived in the early 2009. However, not all future possible applications of RRAM are related to something utopian. The holy grail of computer memory industry—the universal memory—might also be one where the RRAM triumphs, considering its characteristics where both rapid switching and high-density non-volatility storage mechanism are observed.

The plethora of hypothetical as well as actual applications of RRAM does not ensure that their realization and commercialization will go without any obstacles. In fact, the very notion that RRAM currently performs well only in the laboratory, but not as good as in the industry, has driven other competing technologies to develop their product more persistently.

An example is given for the case of HP's progress in their memristor research[91]. In 2007, HP was successful in producing the crossbar switches—the building blocks of the HP's RRAM design—that were more than 20 times denser than Flash memory. In less than a year, Flash memory has upped their density fourfold to eightfold by going to 2- and 3-bits per cell configurations. That unforeseen leap has narrowed HP's competitive advantage. HP's RRAM was then claimed to be only three times the density of flash, evoking memories of the same scenario that has doomed other next-generation memory technologies.

Such unprecedented development was certainly one that is not desired but must be anticipated by the engineers and scientists working on RRAM. In other words, if this technology were to realize its prophecy, it must have been at least accompanied by aggressive R&D activities as well as sufficient capital investments; both of which unfortunately depend on the human factor.

Appendices

Appendix 1: List of important IP (patents) related to RRAM

Patents No.	Assignee(s)	Content	Category
6,204,139	University of Houston	Methods for switching the properties of perovskite materials used in thin film resistors	3
6,473,332	University of Houston	Electrically variable multi-state resistance computing	1
6,531,371	Sharp Laboratories of America, Inc	Electrically programmable resistance cross point memory	4
6,583,003	Sharp Laboratories of America, Inc	Methods of fabricating 1T1R resistive memory array	2
6,586,965	Hewlett Packard Development Company	Molecular crossbar latch	4
6,664,117	Sharp Laboratories of America, Inc	Method for resistance memory metal oxide thin film deposition	2
6,673,691	Sharp Laboratories of America, Inc	Method for resistance switch using short electric pulses	3
6,746,910	Sharp Laboratories of America, Inc	Method of fabricating self-aligned cross-point memory array	1
6,759,249	Sharp Laboratories of America, Inc	Device and method for reversible resistance change induced by electric pulses in non-crystalline perovskite unipolar programmable memory	2
6,762,481	University of Houston	Electrically programmable nonvolatile variable capacitor	4
6,774,054	Sharp Laboratories of America, Inc	High Temperature Annealing of Spin Coated PCMO Thin Film for RRAM Application	2
6,824,814	Sharp Laboratories of America, Inc	Preparation of LCPMO Thin Films which have reversible resistance change properties	2
6,841,833	Sharp laboratories of America, Inc	1T1R resistive memory	1
6,849,891	Sharp laboratories of America, Inc	RRAM memory cell electrodes	1
6,856,536	Unity Semiconductor	Nonvolatile memory with a single transistor and resistive memory element	1
6,868,025	Sharp Laboratories of America, Inc	Temperature compensated RRAM circuit	1
6,888,745	Sharp Kabushiki, Japan	Nonvolatile Memory Device	1

6,906,939	Unity Semiconductor	Rewritable memory with multiple memory layers	4
6,909,632	Unity Semiconductor	Multiple modes of operation in a cross point array	4
6,911,361	Sharp Laboratories of America, Inc	Low temperature processing of PCMO thin film on Ir substrate	2
6,917,539	Unity Semiconductor	High density NVRAM	1
6,937,505	Sharp Kabushiki, Japan	Nonvolatile memory cell and nonvolatile semiconductor memory device	1
6,939,724	Sharp Laboratories of America, Inc	Method for obtaining reversible resistance switches on a PCMO thin film when integrated with a highly crystallized seed layer	3
6,946,702	Winbond Electronics Corp, Taiwan	Resistance random access memory	1
6,955,992	Sharp Laboratories of America, Inc	One mask Pt/PCMO/Pt stack etching process for RRAM applications	2
6,965,137	Unity Semiconductor	Multi-layer conductive memory device	1
6,967,884	Sharp laboratories of America, Inc	RRAM circuit with temperature compensation	1
6,992,920	Sharp Kabushiki, Japan	Nonvolatile Semiconductor Memory Device, and Programming method and erasing method thereof	1
7,002,837	Sharp Kabushiki, Japan	Nonvolatile Semiconductor Memory Device	1
7,009,278	Sharp laboratories of America, Inc	3D RRAM	1
7,016,222	Sharp Kabushiki, Japan	Non volatile semiconductor memory device	1
7,020,012	Unity Semiconductor	Cross point array using distinct voltages	4
7,027,322	Sharp Kabushiki, Japan	EPIR device and semiconductor device utilizing the same	1
7,029,982	Sharp Laboratories of America, Inc	Method of affecting RRAM characteristics by doping PCMO thin films	2
7,038,935	Unity Semiconductor	2-terminal trapped charge memory device with voltage switchable multi-level resistance	4
7,057,914	Unity Semiconductor	Cross point memory array with fast access time	4
7,057,922	Sharp Kabushiki, Japan	Nonvolatile semiconductor memory device	1
7,060,586	Sharp Laboratories of America, Inc	PCMO thin film with resistance random access memory (RRAM) characteristics	2
7,067,862	Unity Semiconductor	Conductive memory device with conductive oxide electrodes	1
7,082,052	Unity Semiconductor	Multi-resistive state element with reactive metal	4

7,084,691	Sharp laboratories of America, Inc	Mono-polarity switchable PCMO resistor trimmer	1
7,095,643	Unity Semiconductor	Rewritable memory with multiple memory layers	1
7,095,644	Unity Semiconductor	Conductive memory array having page mode and burst mode read capability	4
7,098,043	Sharp Laboratories of America, Inc	PCMO spin-coat deposition	2
7,145,791	Sharp Kabushiki, Japan	Memory device having variable resistive memory element	1
7,157,287	Sharp Laboratories of America, Inc	Method of substrate surface treatment for RRAM thin film deposition	2
7,167,387	Matsushita Electric Industrial Co Ltd, Japan	Variable Resistance Element, Method of Manufacturing the Element, Memory containing the element, and method of driving the memory	2
7,180,772	Unity Semiconductor	High density NVRAM	1
7,186,569	Unity Semiconductor	Conductive memory stack with side walls	4
7,186,658	Winbond Electronics Corp, Taiwan	Method and resulting structure for PCMO film to obtain etching rate and mask by inductively coupled plasma	2
7,203,789	Hewlett Packard Development Company	Architecture and methods for computing with reconfigurable resistor crossbars	4
7,235,407	Sharp Laboratories of America, Inc	System and method for forming a bipolar switching PCMO film	2
7,236,389	Sharp Laboratories of America, Inc	Cross-point RRAM memory array having low bit line crosstalk	1
7,309,616	Unity Semiconductor	Laser Annealing of Complex Metal Oxides (CMO) Memory Materials for Non-volatile Memory Integrated Circuits	2
7,326,979	Unity Semiconductor	Resistive memory device with a treated interface	2
7,339,813	Sharp Laboratories of America, Inc	Complementary output resistive memory cell	4
7,359,888	Hewlett Packard Development Company	Molecular-junction-nanowire-crossbar-based neural network	4
7,363,604	SanDisk 3D LLC	Apparatus and method for programming an array of nonvolatile memory cells including switchable memory elements	3
7,372,718	Sony Corporation	Storage and semiconductor device	1
7,372,753	Unity Semiconductor	Two-cycle sensing in a two-terminal memory array having leakage current	3

73

7,433,222	Sharp Kabushiki, Japan, Institute of Adv. Ind. Sci & Tech	Nonvolatile Semiconductor Memory Device	1
7,447,828	Blaise Mouttet	Programmable crossbar signal processor used as morphware	4
11/301,869*	B.C. Stipe	Unipolar RRAM device and vertically stacked architecture	1
11/510,428*	Sharp Laboratories of America, Inc	Fabrication of a high speed RRAM having narrow pulse width programming capabilities	2
11/740,309*	University of Pennsylvania	Nonvolatile resistance switching oxide thin film devices	2
12/026,701*	Micron Technology, Inc	Resistive Memory Device	1

Note:

   * Patent still pending

Category: 1-- Memory architecture and design; 2--Fabrication method; 3--Testing & characterization; 4--Others development

Appendix 2: Detailed Cost Model Spreadsheets

RRAM-based SSD - COST SUMMARY			
MIT - Materials Systems Laboratory			

VARIABLE COSTS	per wafer	per year	percent
Material Cost	$71.26	$35,631,243	53.09%
Energy Cost	$0.02	$11,916	0.02%
Labor Cost	$10.50	$5,250,000	7.82%
Total Variable Cost	$81.79	$40,893,159	60.93%

FIXED COSTS	per piece	per year	percent	investment
Lithography Machine Cost	$28.69	$14,346,149	21.38%	$72,000,000
RIE Equipment Cost	$0.64	$318,803	0.48%	$1,600,000
Furnace	$0.56	$278,953	0.42%	$1,400,000
CMP	$0.35	$175,342	0.26%	$880,000
CVD	$1.28	$637,607	0.95%	$3,200,000
Sputter	$1.91	$956,410	1.43%	$4,800,000
Spin Coat	$0.33	$199,252	0.30%	$1,000,000
Fixed Cleanroom Cost	$0.32	$159,402	0.24%	$800,000
Building Cost	$2.00	$1,000,000	1.49%	$1,000,000
Lithography Mask Cost	$13.40	$6,700,000	9.98%	$6,700,000
Maintenance Cost	$0.20	$98,000	0.15%	-
Installation Cost	$2.71	$1,353,001	2.02%	$6,790,400
Total Fixed Cost	$52.38	$26,222,918	39.07%	$85,880,000

| Total Fabrication Cost | $134.17 | $67,116,077 | 100.00% | |

INPUT VARIABLE

INPUT

Production Volume	500,000.00	cells
Production Capacity	600,000.00	units

EXOGENOUS DATA

Annual Production Volume	500,000.00	(cells/yr)
Facility Production Capacity	600,000.00	(cells/yr)
Working Days/Yr Equiptment	365.00	days
Unplanned Downtime	0.96	hrs/day
No (i.e., none NOT No.) Operations	0.00	hrs/day
Planned Paid	0.00	hrs/day
Planned Unpaid	0.00	hrs/day
Price, Building Space	1,000,000.00	total for building/yr
Price of cleanroom	800,000.00	total for building
Price of Electricity	0.05	/kWh
Accounting Life of Machine	10.00	yrs
Discount Rate	0.15	
No of die can be made in one wafer	241	die
No of wafer in 1 lot	40	wafer

I. Scrap Reject Rate

Si-wafer	0.01
Lithography	0.04
Etching	0.04
Furnace	0.01
CMP	0.01
CVD	0.01
Sputter	0.01
Spin-coating	0.04

II. Equipment Cost

Lithography Eqt Cost	$18,000,000	(1 unit)
Etching	$400,000	(1 unit)
Furnace	$350,000	(1 unit)
CMP	$220,000	(1 unit)
CVD	$800,000	(4 unit)
Sputter	$1,200,000	(4 unit)
Spin-coating	$250,000	(1 unit)

III. Logistics and Location

Lines	4	
Floor space	120.00	m2
Overhead for cleanroom	$98,000	/year

IV. Energy Consumption

Energy Consumption lithography	0.90	kWh
Energy Consumption RIE	1.50	kWh
Energy Consumption for Furnace	2.67	kWh
Energy Consumption for CMP	0.70	kWh
Energy Consumption for CVD	1.20	kWh
Energy Consumption for Sputtering	1.50	kWh
Energy Consumption for Spin Coating	0.50	kWh
Energy Consumption for Other	4.00	kWh

V. Labors

Direct Laborers Per Lithography	0.50	person/eqpt
Direct Laborers Per RIE	0.50	person/eqpt
Direct Laborer Per Furnace	0.50	person/eqpt
Direct Laborers Per CMP	0.50	person/eqpt
Direct Laborers Per CVD	0.50	person/eqpt
Direct Laborer Per Sputtering	0.00	person/eqpt
Direct Laborer Per Spin Coating	0.00	person/eqpt
Direct Supervisor per Process	1.00	person/eqpt
Direct Wages Technician (w/ benefits)	$20	/hr
Direct Wages Supervisor (w/ benefits)	$30	/hr

Process Calculation

Actual Production Capacity	600,000	
Effective Capacity	735,154	
Capacity and Volume are Consistent?	TRUE	
Available Operating Time Lithography	8,410	hours / year
Available Operating Time RIE	8,410	hours / year
Available Operating Time RCA	8,410	hours/year
Effective Cycle Time presses total	1.50	hr / lot
Required Operating Time (produced)	4,688	hours / year
Required Operating Time (capacity)	6,892	hours / year
Run-Time for One Machine (produced)	56%	
Run-Time for One Machine (capacity)	82%	
Number of Parallel Streams of Presses	4	
Annual Energy Consumption	60,796.88	kWh

Cycle Time Calculations

Lithography	1.50	hr/lot
RIE	0.50	hr/lot
Furnace	1.33	hr/lot
CMP	0.50	hr/lot
CVD	0.80	hr/lot
Sputtering	0.80	hr/lot
Spin-Coating	0.20	hr/lot
Total Cycle Time (hour/lot)	1.50	hr/lot

Materials Requirements:

Silicon Wafer

Unit price 12"-Si wafer	$225.00	
Annual Wafer Required	133,004	wafer/year
Amount spent on Si wafer	$29,925,903	/year

Lithography Masks

Ann. Lithography Masks Required	67.00	mask/yr
Unit price of the mask	$100,000.00	$/mask
Amount spent on Masks	$6,700,000.00	/year

Photoresist

Unit price of the photoresist	$20.000	$/wafer
Amount spent on photoresist	$2,660,080.228	/year

Perovskite

Unit price of the peroskite	$1.310	$/wafer
Amount spent on Perovskite film	$174,235.255	/year

TiN

Unit price of the bulk TiN	$2.700	$/wafer
Amount spent on dielectric gate (TiN)	$359,110.831	/ year

Hf metal

Unit price per wafer	$7.600	$/wafer
Amount spent on metal gate (HfO2)	$1,010,830.487	/ year

TaN

Unit price per wafer	$0.166	$/wafer
Amount spent on TaN diffusion barrier	$22,078.666	/year

Pt electrode

Unit price per wafer	$4.320	$/wafer
Amount spent on Pt electrode	$574,577.329	/year

Cu metallization

Unit price per wafer	$4.000	$/wafer
Amount spent on Cu	$532,016.046	/year

SiO2

Unit price per SiO2	$2.800	$/wafer
Amount spent on SiO2	$372,411.232	/year

References

[1]   Executive Summary, ITRS 2007 Report. Article retrieved from http://itrs.net on Feb 2009

[2]   M. R. Wordeman, *Concise encyclopedia of computer science*, p 496, Wiley (2004)

[3]   G. F. Derbenwick, J. E. Brewer (ed), *Nonvolatile Memory Technologies with Emphasis on Flash:* p619-623, IEEE Press (2008)

[4]   Lithography, ITRS 2007 Report. Article retrieved from http://itrs.net at Feb 12, 2009

[5]   G. I. Meijer, *Science*, **319**, 1625 (2008)

[6]   Q. Guo, *Evaluation on the thin-film phase change material-based technologies*, M. Eng thesis, DMSE-MIT (2006)

[7]   M. N. Kozicki, *et al, 6th Non-Volatile Memory Technology Symposium, IEEE*, 89 (2005)

[8]   S. Q. Liu, N. J. Wu, and A. Ignatiev, *Appl. Phys. Lett*, **76**, 2749 (2000)

[9]   S. Seo *et al, App. Phys. Lett.* **85**, 5655 (2004)

[10]   B. J. Choi *et al, J. Appl. Phys.* **98**, 033715 (2005)

[11]   S. F. Karg *et al, IBM J. Res and Dev,* **52**, 481 (2008)

[12]   A. Sawa et al, *Jpn. J. Appl. Phys.,* **44**, L1241 (2005)

[13]   L. Ma *et al, Appl. Phys. Lett.* **82**, 1419 (2003)

[14]   B. Mukherjee and A. J. Pal, *Appl. Phys. Lett.* **85**, 2116 (2004)

[15]   T. Oka and N. Nagaosa, *Phys. Rev. Lett.* **95**, 26603 (2005)

[16]   K. Nagashima *et al, Appl. Phys. Lett.* **94**, 242902 (2009)

[17]   A. Beck *et al, Appl. Phys. Lett.* **77**, 139 (2000)

[18]   Y. Watanabe *et al, App. Phys. Lett.* **78**, 3738 (2001)

[19]   A. Ignatiev *et al, Phys. Stat. Sol. (B)*, **243**, 2089 (2006)

[20]   N. A. Tulina *et al, Physica C*, **385**, 563 (2003)

[21]   S. Tsui *et al, Appl. Phys. Lett.* **83**, 957 (2003)

[22]   R. Fors, S. I. Khartsev and A. M. Grishin, *Phys. Rev. B*, **71**, 045305 (2005)

[23]   A. Odagawa *et al, Phys. Rev. B*, **70**, 224403 (2004)

[24]   Ch. Jooss *et al, Proc. Nat. Acad. Sci.* **104**, 13597 (2007)

[25]   I. Bersuker, *The Jahn-Teller effect,* Cambridge University Press (2006)

[26]   Ch. Jooss *et al, Phys. Rev. B,* **77**, 132409 (2008)

[27]   W. Pan and D. R. Evans, *MRS Symp. Proc.* **933**, G04-05 (2006)

[28]   L. L. Hench and J. K. West, *Principles of Electronic Ceramics*, John Wiley (1990).

[29]   A. Sawa *et al, Appl. Phys. Lett.* **85**, 4073 (2004)

[30]   S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed (2006)

[31]   M. J. Rozenberg, I. H. Inoue, and M. J. Sánchez, *Phys. Rev. Lett.*, **92**, 178302 (2004)

[32]   C. A. Marianetti, G. Kotliar and G. Ceder, *Nat. Mater.* **3**, 627 (2004)

[33]   M. Imada, A. Fujimori, and Y. Tokura, *Rev. Mod. Phys.,* **70**, 1039 (1998)

[34]   I. H. Inoue *et al, 6th Non-Volatile Memory Technology Symposium IEEE*, **10**, 131 (2005)

[35]   B. E. Saleh and M. C. Teich, *Fundamentals of Photonics*, 2nd ed (2007)

[36]   J. G. Simmons and R. R. Verderber, *Proc. Royal Soc. London Ser. A*, **301**, 77 (1967)

[37]   M. J. Rozenberg, I. H. Inoue and M. J. Sánchez, *Appl. Phys. Lett.*, **88**, 033510 (2006)

[38]   H. Mori *et al, J. Am. Chem. Soc.* **124**, 1251 (2002)

[39]   M. J. Rozenberg, I. H. Inoue, and M. J. Sánchez, *Thin Solid Films*, **486**, 24 (2005)

[40]   M. J. Sánchez, M. J. Rozenberg, and I. H. Inoue, *App. Phys. Lett.*, **91**, 252101 (2007)

[41]   D. S. Jeong, H. Schroeder and R. Waser, *Electrochem. Solid-State. Lett.,* **10**, G51 (2007)

[42]   Y. Hosoi *et al*, *Tech. Dig., – Int. Electron Devices Meetings*, 793 (2006)

[43]   R. Waser and M. Aono, *Nature Mat.* **6**, 833 (2007)

[44]   I. G. Back, et al, *IEDM Paper Technical Digest*, p587 (2004)

[45]   R. Waser, *Microelec. Engin.* **86**, 1925 (2009)

[46]   S. Tsui et al, *Appl. Phys. Lett.* **89**, 123502 (2006)

[47]   Y. B. Nian, et al, *Phys. Rev. Lett.* **98**, 146403 (2007)

[48]   Muller et al, *29th European Solid-State Circuit Conference*, ESSCIRC (2004)

[49]   R. F. Freitas, W. W. Wilcke, *IBM J. Res. & Dev.,* **52**, 439 (2008)

[50]   W. W. Zhuang *et al*, *Elec. Devices Meeting 2002 Digest Int, IEEE,*193 (2002)

[51]   U. S. Patent No. 6,841,833

[52]   U. S. Patent No. 7,016,222

[53]   U. S. Patent No. 6,946,702

[54]   U. S. Patent Application No. 2004/0202041 A1

[55]   U. S. Patent No. 6,372,659

[56]   U. S. Patent No. 6,291,866

[57]   T. Dao *et al*, *Integ. Circ. Des. Tech. and Tut. 2008, IEEE Intl Conf on,* 109 (2008)

[58]   M. Lemberger, A. Baunemann, and A. J. Bauer, *Microelec. Reliability,* **47**, 635 (2007)

[59]   T. Yamada *et al*, *Microelec. Reliability,* **41**, 697 (2001)

[60]   T. Aaltonen *et al*, *Chem. Mater.,* **15**, 1924 (2003)

[61]   U. S. Patent No. 6,962,648

[62]   U. S. Patent No. 7,098,043

[63]   U. S. Patent No. 6,911,361

[64]   L. O. Chua, *IEEE Transactions on Circuit Theory,* **18**, 507 (1971)

[65]   J. M. Tour, H. Tao, *Nature* **453**, 42 (2008)

[66]   R. Stanley Williams, *Nature* **453**, 80 (2008)

[67]   L. O. Chua and S. M. Kang, *Proc. IEEE.,* **64**, 209 (1976)

[68]   M. Sapoff and R. M. Oppehneim, *Proc. IEEE,* **51**, 1292 (1963)

[69]   Y. V. Pershin, *et al*, *Nature Precedings*, hdl:10101/npre.2008.2431.1

[70]   M. di Ventra, Y. V. Pershin, L. O. Chua, arXiv:0901.3682v1 [cond-mat.mes-hall]

[71]   Y. Kim et al, *Appl. Phys. Lett.,* **78**, 934 (2001)

[72]   S. Q. Liu, N. Wu, A. Ignatiev, J. Li, *J. Appl Phys.* **100**, 056101 (2006)

[73]   Semiconductor Industry Assoc., *Reports on Global Semiconductor Chip Sales*, (2007)

[74]   G. S. Kim, J. D. Ha, *Semiconductor Sector Report*, Goodmorning Securities, Feb 2009

[75]   FirstGlobal, U.S. Research on Technology (Memory) Sector, (2009)

[76]   Samsung Technological Feature on HDD and SSD, 2008

[77]   DRAMeXchange, *Weekly Research Market View on Computex 2009.* Article retrieved from http://www.dramexchange.com/WeeklyResearch/PrinterFriendly/2060.html on Jun 10, 2009

[78]   Paisit, *Solid State Disk & Storage Performance,* Sun Microsystems. Article retrieved from http://blogs.sun.com/paisit/entry/ solid_state_disk_ssd_and on April 12, 2009

[79]   Web-Feet Research, *Solid State Drives (SSD) Markets and Application (Annual),* (2008)

[80]   IDC, Research report on HDD Industry, (2008)

[81]   DRAMeXchange, Market report on Flash memory, (2009)

[82]   Frost & Sullivan, Semiconductor IP-Market, (2006)

[83] Z. Kerekes (ed), *Solid State Disks Buyers Guide*. Article retrieved from http://www.storagesearch.com/ssd-buyers-guide.html on June 1, 2009

[84] D. K. de Vries, *Semiconductor Manufacturing, IEEE Transactions on*, **18**, 136, (2005)

[85] Applied Materials, website: http://appliedmaterials.com; B. Matas, *Semiconductor Market Forecast 2009-2010*, GSA's EDA/Design & IP Advisory Panels, IC Insights, Feb 2009; S. Hector, Mask Magnification Working Group, International Sematech (2004)

[86] Sigma-Aldrich, website: http://sigmaaldrich.com/united-states.html; Strem Chemicals, website: http://www.strem.com/catalog/

[87] J. Ng, *Intel Cuts SSD Prices*, DailyTech. Article retrieved from http://www.dailytech.com/Intel+Cuts+SSD+Prices+Makes+More+18Inch+SSDs+for+Netbook s+and+Small+Laptops/article14992.htm on May 6, 2009

[88] P. Seidel, V. Bakshi (ed), *EUV Lithography, Ch. 12*, p585, Wiley (2008)

[89] H. Tsuchikawa, M. Takakuwa, S. Sugatani, *4th ISMI Symposium on Manufacturing Effectiveness*, October 2007; B.J. Lin, *Microelectronic Engineering*, **83**, 604 (2006); D2S, Inc. http://www.direct2silicon.com/read/optical_lithography_related

[90] D2S and E-beam Initiative website. Article retrieved from http://www.direct2silicon.com/dfeb on May 10, 2009

[91] R.C. Johnson, *Will memristors prove irresistible?* EETimes.com. Article retrieved from: http://www.eetimes.com/showArticle.jhtml?articleID= 210004310 on Jan 20, 2009

[92] S. Bush, *HP nano device implements memristor, Electronics Weekly*. Article retrieved from http://electronicsweekly.com/Articles/2008/05/02/43658/hp-nano-device-implement s-memristor.htm on Feb 12, 2009

[93] Unity Semiconductor Press Release. Article retrieved from http://www.unitysemi.com/ downloads/ news/Unity_Launch_release.pdf., on June 20, 2009

[94] R. Meyer, et al, *9th Non-volatile Memory Technology Symposium, IEEE*, 54 (2008)

[95] Unity Semiconductor Press Release. Article retrieved from http://www.unitysemi.com/ downloads/ news/Unity_Funding_Release.pdf on June 20, 2009

[96] M. LaPedus, *Unity rolls storage-class memory*, EETimes.com. Article retrieved from http://www.eetimes.com/showArticle.jhtml?articleID= 217500734 on June 20, 2009

[97] L. Wu, et al, *Appl. Phys. Lett.* **94**, 243115 (2009)

[98] B. Mouttet, *Unity Semiconductor's RRAM patents*. Article retrieved from http://tinytechip.blogspot.com/2008/07/unity-semiconductors-rram-patents.html (personal blog) on May 10, 2009

[99] M. LaPedus, *Australian startup claims RRAM breakthrough*, EETimes.com. Article retrieved from: http://www.eetimes.com/showArticle.jhtml?articleID=213400587 on April 12, 2009

[100] P. Clarke, *IMEC begins work on resistive RAM*, EETimes.com. Article retrieved from: http://www.eetimes.com/showArticle.jhtml?articleID=211200324 on April 12, 2009

[101] N. A. Tulina, *Physics – Uspekhi*, **50**, 1171 (2007)

[102] R. Münstermann, *et al, Non-Volatile Memory Technology Symposium 2008*, Pacific Grove, California, 12.11.2008 - 16.11.2008

[103] C. Cen et al., *Nat. Mat.*,**7**, 298 (2008)

[104] Y. Zhao, *Wafer Level Reliability of Adv. CMOS Devices*, Nova Science (2008)

[105] Micron, Inc. MLC vs SLC NAND Flash. Article retrieved from: http://www.micron.com/products/nand/mlc-slc on June 12, 2009