Design and Optimization of TMO-ReRAM Based Synaptic Devices

J.F. Kang\textsuperscript{1#}, B. Gao\textsuperscript{1}, P. Huang\textsuperscript{1}, Z. Chen\textsuperscript{1}, L.F. Liu\textsuperscript{1}, X.Y. Liu\textsuperscript{1}, S.M. Yu\textsuperscript{2}, H.-S. P. Wong\textsuperscript{3}

#E-mail: kangjf@pku.edu.cn

\textsuperscript{1}Institute of Microelectronics, Peking University

\textsuperscript{2}School of CIDSE, Arizona State University

\textsuperscript{3}Department of Electrical Engineering, Stanford University
Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
Introduction

Resistive Switching (RS)

- Many materials have been used to demonstrate the reversible bi-stable resistance states (LRS and HRS), which can be switched by voltage, named as resistive switching (RS)

RRAM

- These RS materials can be used to construct a device, with a typical sandwiched structure, termed as RRAM (Resistive-switching Random Access Memory).
Introduction

Two Switching Modes [#]

Unipolar

Bipolar

depend on amplitude of applied voltage but not on polarity

depend on the polarity of the applied voltage

Introduction

- Excellent performances have been demonstrated in transition metal oxide (TMO)-ReRAM [1-6].
  - **Scalability:** <10nm devices demonstrated [1-2]
  - **Compatibility with CMOS using fab-friendly materials** [1-4]
    - HfO₂, TaOₓ, WOₓ, Ti, Ta, TiN, NiSi
  - **Switching speed:** <1ns [6]
  - **Switching voltage:** <1.5V
  - **Endurance:** >10¹⁰ cycles [5]
  - **Retention:** >10 yrs [6]
  - **Read disturb:** >10¹⁰ times [3]

Introduction

 Capability to High Density Integration [1,2]

 Vertical MOSFET

✓ 32/16 Gb Test Chips have been demonstrated [3,4]

Introduction

New Function Application

- Concept of RRAM based memristor [1]
  - Memristive switches: both store logic values and perform logic operations [2]

Introduction

- RRAM based synapses for neuromorphic computing systems [#]

[...] S.M. Yu et al, IEDM2012, p.239 (Stanford and PKU)

Most demonstrated in the bipolar switching mode
For applications

- Understand the physical mechanisms of RS
- Seek technical solutions to construct RRAM devices to achieve targeted performances [1]

In this talk, we will also address

- Low energy and robust synapse performances of TMO-RRAM [2, 3]
- Potential for application in a neuromorphic visual system [2]

Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
For the resistive switching (RS) behavior of TMO-RRAM

- Filament effect has been widely accepted
  - RS is due to the formation and rupture of conducting filaments

However, the physical natures of filaments and the crucial effects to dominate the formation and rupture of filaments are still argued

- Conducting filament (CF) type: Vo or metallic ions?
- Dominant effect for SET/RESET: G-R or S/D? Thermal or E-field?
- Mechanisms of unipolar and bipolar: Same or not?
A Unified Physical Mechanism [1,2]

- To clarify fundamental properties of resistive switching behaviors in TMO-ReRAM


The mechanism is based on filament effect on RS [3]

Unified Physical Mechanism

The unified physical mechanism is proposed to clarify these argued issues:

- Microscopic physical properties correlated with resistive switching in TMO-based RRAM (including unipolar and bipolar)

  - To explain various resistive switching characteristics observed in TMO-RRAM
  - To predict performances of TMO-RRAM
Unified Physical Mechanism

Schematic microscopic properties of RS in TMO-RRAM
(B. Gao et al, IEDM2011, p.417)

1. Filament: A percolation path consisting of $V_0$ defects
2. Formation and rupture of filaments are correlated with generation and recombination of $V_0$
3. Forming/SET: Generation of new $V_0$ defects and $O^{2-}$ ions induced by E-field and thermal effects in rupture region

$V_0$ defects may be in different states:
- Filled state ($V_0$) with 2 electrons in $V_0$
- Unfilled state ($V_0^{2+}$) w/o electron in $V_0$

$p = \exp\left[\left(\frac{eLE - \epsilon_f}{kT}\right)\right]$
Unified Physical Mechanism

Schematic microscopic properties of RS in TMO-RRAM

4. **RESET:** Recombination among charged $V_0^{2+}$ and $O^{2-}$

\[
V_O^{E-Field} \rightarrow V_O^{2+} + 2e^- \\
V_O^{2+} + O^{2-} \rightarrow LO
\]

5. **Two essential conditions for RESET**
   1) Occurrence of $V_0^{2+}$ states induced by a critical E-field
   2) Presence of moveable $O^{2-}$

Formation of the state $V_0^{2+}$ in the filament at a critical E-field

- significant capture section
- stable recombination state (LO)
6. Conduction Properties: due to electron transport along Vo filaments

- Semiconductor-like: $V_o$ are separated from each other
- Metallic-like: $V_o$ are closed each other in the clustered
- First principle calculations support this opinion
Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
The resistive switching characteristics are correlated with geometry of Vo filament generation, recombination, and distributions of Vo.

It is crucial to control Vo distributions and filament geometry to achieve targeted performances.
Defect Engineering Approach

According to crystal defect theory, the generation and recombination probability of Vo is governed by

\[ p = \exp\left(\frac{\gamma E_{loc} - \varepsilon_a}{kT}\right) \]

\( E_{loc} \): Local electric field

\( \varepsilon_a \): Formation energy of Vo

A Defect Engineering Approach is proposed [*]

A Defect Engineering Approach is proposed

**A. Material-Oriented Cell Design**

A Defect Engineering Approach is proposed

**B. Innovation Operation Scheme**
## A. Material-Oriented Cell Design

Calculated formation energy $\varepsilon_a$ of V$_o$. [1, 2]

<table>
<thead>
<tr>
<th></th>
<th>Undoped (eV)</th>
<th>Ti (eV)</th>
<th>Al (eV)</th>
<th>La (eV)</th>
<th>Ga (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HfO$_2$</td>
<td>6.53/6.40$^a$</td>
<td>6.48</td>
<td>4.09</td>
<td>3.42</td>
<td>–</td>
</tr>
<tr>
<td>ZrO$_2$</td>
<td>6.37/6.09$^b$</td>
<td>6.11</td>
<td>3.66</td>
<td>3.74</td>
<td>3.77</td>
</tr>
</tbody>
</table>

$^a$ A. S. Foster et al. PRB 65, 174117(2002); $^b$ A. S. Foster et al. PRB 64, 224108(2001); $^c$ T. R. Paudel et al. PRB 77, 205202(2008)

Trivalent La or Al doping could effectively reduce $\varepsilon_a$

A. Material-Oriented Cell Design

- In the resistive switching (RS) layers of Al- or La-doped HfO$_2$ or ZrO$_2$ [1-2]
  - $V_0$ are preferentially generated near the trivalent Al or La sites
  - Filaments are preferentially formed along the dopant sites
  - Better controllability of resistive switching could be achieved by using proper doping approaches

Defect Engineering Approach

A. Material-Oriented Cell Design: Doping Effect

Vo distributions and CFs are full-randomly

Vo and CFs are formed near the dopant sites
Improved Uniformity by proper doping

Expected uniformity improvement is identified by experiments
- Better controllability on RS processes achieved in doped HfOx devices
- This is beneficial for RRAM as a synapse
Defect Engineering Approach

B. Innovation Operation Scheme

\[ N(Vo) = \nu t \exp[ql \left( \frac{E_{\text{loc}} - \varepsilon_a}{kT} \right) / kT] \]

- Vo density is dependent on local electric field and switching time
- Operation schemes (switching time and local electric field) can be used to control Vo distributions

Different operation schemes can be expected to achieve different response characteristics!!
Non linear resistance change as a function of pulses is observed when short pulses are applied.
B. Innovation operation scheme

- **Defect Engineering Approach**

- **Nearly linear resistance change with pulses is realized when wider pulses are applied.**
Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
Neuromorphic Visual Systems

- A great amounts of synapses are needed
- A typical CMOS-based binary synapse consisted of a 8T-SRAM cell [*]
- New synapse is needed

[*] SRAM based CMOS hardware (IBM, CICC 2011)
TMO-based Synaptic Devices

- TMO-RRAM-based synapse is promising

Analogy between biological and artificial RRAM synapse.

Analogy between biological and RRAM based neural networks.
Artificial Visual System-based on RRAM and Winner-Take-All algorithm is constructed.

Optimization of Synapse

1st layer: 32×32 neurons; 2nd layer: 4×4 neurons
between 1st layer and 2nd layer: 16348 RRAM synapses
Multi-level resistance states and ultra-low spike energy <1pJ are demonstrated [#]

Dependence of energy/spike on initial R for training process

Measured and fitted training process with pulse amplitudes.

[#] S. Yu, et al, IEDM2012, p.239
A model is developed for the training process of TMO-RRAM synapse

- Resistance variation effect during training process
- Model parameters can be extracted from measured data.
Resistance Evolution under 400 RESET pulses

- In low resistance regime, fluctuation is smaller but suffers from high spike energy.
- In high resistance regime, low spike energy but larger fluctuation presented.

Larger fluctuation or variation may cause degradation of recognition accuracy of the neuromorphic systems.
Training Images and Initial Conductance Map

2D Gaussian bar:
Random center and random orientation

Before the training randomized around 20kΩ
Simulated System Performances

Resistance Diverges and Orientation Map Emerges

During the training

After the training

2015 SPICE Workshop
June 29-July 3 2015
Mainz, Germany
Optimized Synaptic Devices

Can we realize synaptic performances with both low spike energy and high recognition accuracy? ?

Geometric mean of more than 2 devices in parallel can significantly suppress the impact of intrinsic fluctuation effect.
Optimized architecture of a neuromorphic system using robust synapse is proposed

- A 1D1R synaptic cell is introduced
- 1D is applied to perform logarithm function on the device resistance

Geometric mean calculation on resistance is replaced by the logarithm function.
Simulated System Accuracy

- Single RRAM device
- Geometric mean of two devices
- Two parallel 1D-1R cells

- Significant improvement on recognition accuracy is achieved by the architecture of 2 parallel 1D1R.
- Array integration approach is a great challenge
- 3D vertical ReRAM array architecture as synapses
  - Easily to achieve high density of integration
  - Significantly to immunize resistance variation during training process of synapses

A synapse: devices in the same pillar electrode
3D Vertical RRAM Arrays

- Measured training process of top and bottom ReRAM devices in the 3D vertical array
  - 2 layered devices are fabricated
  - Nearly constant device performance both in top and bottom layers is measured.
  - Significantly improved accuracy achieved.
TMO-based Synaptic Devices

- Measured training process for the 3D vertical synaptic devices
  - Different initial R states can be achieved by different current compliances
  - Initial R is set to $\sim 1\,\text{M}\Omega$, maximum energy consumption per spike $< 1\,\text{pJ}$.

Outline

- Introduction
- Physical Mechanism
- Defect Engineering Approach
- Optimization of Synapse
- Summary
Summary

- A unified physical mechanism is proposed to elucidate the resistive switching of TMO-RRAM
- A defect engineering approach is developed to design and optimize RRAM performances
- Excellent controllability on RS behaviors is demonstrated in optimized RRAM devices based on the defect engineering approach.
Summary

- Multi-level resistance states are realized in the optimized RRAM
- Robust synaptic behaviors with sub-pJ energy per spike are realized in the optimized RRAM
- Optimized architectures of TMO-RRAM synapse are proposed to improve system performances.