Resistive switching phenomena: A review of statistical physics approaches

Jae Sung Lee,1,a) Shinbuhm Lee,2,a) and Tae Won Noh3,4,b)
1School of Physics, Korea Institute for Advanced Study, Seoul 130-722, South Korea
2Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
3Center for Correlated Electron Systems, Institute for Basic Science, Seoul 151-747, South Korea
4Department of Physics and Astronomy, Seoul National University, Seoul 151-747, South Korea

(Received 3 May 2015; accepted 19 July 2015; published online 31 August 2015)

Resistive switching (RS) phenomena are reversible changes in the metastable resistance state induced by external electric fields. After discovery ~50 years ago, RS phenomena have attracted great attention due to their potential application in next-generation electrical devices. Considerable research has been performed to understand the physical mechanisms of RS and explore the feasibility and limits of such devices. There have also been several reviews on RS that attempt to explain the microscopic origins of how regions that were originally insulators can change into conductors. However, little attention has been paid to the most important factor in determining resistance: how conducting local regions are interconnected. Here, we provide an overview of the underlying physics behind connectivity changes in highly conductive regions under an electric field. We first classify RS phenomena according to their characteristic current–voltage curves: unipolar, bipolar, and threshold switchings. Second, we outline the microscopic origins of RS in oxides, focusing on the roles of oxygen vacancies: the effect of concentration, the mechanisms of channel formation and rupture, and the driving forces of oxygen vacancies. Third, we review RS studies from the perspective of statistical physics to understand connectivity change in RS phenomena. We discuss percolation model approaches and the theory for the scaling behaviors of numerous transport properties observed in RS. Fourth, we review various switching-type conversion phenomena in RS: bipolar-unipolar, memory-threshold, figure-of-eight, and counter-figure-of-eight conversions. Finally, we review several related technological issues, such as improvement in high resistance fluctuations, sneak-path problems, and multilevel switching problems. © 2015 AIP Publishing LLC.

[http://dx.doi.org/10.1063/1.4929512]

TABLE OF CONTENTS

I. INTRODUCTION .................................. 2
A. A brief history of resistive switching (RS) ........ 2
B. The importance of inhomogeneity and switching-type conversions 3
C. Scope and outline of this review .................. 5

II. RESISTIVE SWITCHING AND ITS CLASSIFICATION ............ 5
A. Basic operation of electric field-induced resistive switching .......... 5
B. Classification of resistive switching phenomena ...................... 6
1. Unipolar switching ................................ 7
2. Bipolar switching .................................. 8
3. Threshold switching ................................ 8

III. MICROSCOPIC MECHANISMS OF RESISTIVE SWITCHING IN OXIDES .... 8
A. Migration of oxygen vacancies ...................... 9

IV. IMPORTANCE OF INHOMOGENEITY AND STATISTICAL PHYSICS APPROACHES TO RESISTIVE SWITCHING .......... 22
A. Limitations of microscopic approaches: Fluctuations due to inhomogeneity 22

1. How do oxygen vacancies play a role in resistive switching? ........ 9
2. Evidence of oxygen vacancy migration .................. 9
3. Microscopic origins for oxygen vacancy migration .................. 11
B. Unipolar switching .................................. 12
1. What happens in an oxide during unipolar switching? ............ 12
2. Forming and set processes via soft dielectric breakdown .......... 15
3. The reset process via rupturing of conducting filaments .......... 17
C. Bipolar switching .................................. 18
1. What causes polarity-dependent operation? .................... 18
2. Forming process ................................... 19
3. Models for set and reset processes ....................... 20

---

a)J. S. Lee and S. Lee contributed equally to this work.
b)Electronic mail: twnoh@snu.ac.kr

1931-9401/2015/2(3)/031303/57/$30.00 2, 031303-1 © 2015 AIP Publishing LLC
I. INTRODUCTION

A. A brief history of resistive switching (RS)

RS refers to physical phenomena whereby the resistance of a dielectric material changes in response to the application of a strong external electric field. It differs from dielectric breakdown phenomena, which result in a permanent reduction in resistance (and, in many cases, incurable damage to the sample) such that switching back to the original state is not possible. The RS process is reversible and can be repeated a number of times. Typically, the change in resistance is non-volatile (i.e., the resulting resistance states can be maintained for a long time following the removal of the externally applied electric field). Note that these phenomena occur in numerous insulating materials, including oxides, nitrides, chalcogenides, semiconductors, and organic materials. However, RS phenomena have been studied most widely in oxides; therefore, in this review, we will focus mainly on RS phenomena in oxides.

Studies of RS phenomena can be dated back to the early 1960s. The effect was first reported by Hickmott. Figure 1(a) shows the report by Hickmott on RS using the current–voltage (I–V) relationship for an aluminum oxide film sandwiched between two metal electrodes. The as-grown aluminum oxide film was a good insulator, showing an exponential relationship between current and voltage. As the applied bias increased further, the resistance suddenly dropped; this sudden drop in resistance (originally termed electroforming) corresponds to RS. Extensive studies of RS followed in the 1960s and 1970s. A proposal to use this phenomenon in non-volatile memory devices appeared during this early period and a couple of review articles were published.

In the 1980s and 1990s, research interest in RS declined significantly. This was partly due to the slow progress in understanding the physics of RS and how to control RS phenomena. Furthermore, interest in technological applications of RS was dwarfed by the enormously successful development of silicon-based electronics. However, during this period, there were two important developments in science and technology related to RS. First, following the discovery of high-temperature superconductivity in cuprates, the physics of transition metal oxides, especially in thin film forms, attracted much renewed interest. There was a high demand for high-quality, oxide thin films, leading to significant improvements in oxide thin film fabrication methods, including radio-frequency sputtering and pulsed-laser deposition. There were also significant advances in experimental techniques, including spectroscopic and microscopic methods. Such advancements made it possible to investigate numerous physical properties of oxide thin films in the energy and space domains. Second, the scaling of the size of features in electronic devices accompanied the remarkable development of semiconductor technology. The areal density of devices doubled approximately every 18 months, a phenomenon referred to as Moore's law. In the early 2000s, feature size was reduced to the level of about 100 nm, which led many people to develop serious concerns that the
This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 147.46.19.34 On: Tue, 13 Oct 2015 06:35:15
review of RS for memristive devices can be found in Refs. 38, 48, and 50.

Most of the aforementioned review articles have tried to cover the numerous research efforts related to the microscopic origins of RS. Namely, they mainly focused on how atomic rearrangement or migration can modulate the resistance of a local region inside an RS medium. Such approaches are useful for providing a picture of the microscopic changes that occur in local regions in a sample during RS operations that make originally insulating regions into conductors. However, the articles did not explicitly take into account the detailed interconnecting patterns of conducting regions formed during the RS operation.

Information on these interconnecting patterns is of crucial importance for understanding the electrical properties in RS. This is because total resistance should vary strongly, depending on how the local conducting regions are interconnected, in a way similar to that of inhomogeneous metal-insulator composites. For example, consider an insulating medium (yellow-colored region) with top and bottom electrodes as shown in Fig. 2(a). Suppose that conducting materials (blue-colored region) with a volume fraction \( p \) are added to the insulating medium, as shown in Figs. 2(b)–2(d). The total resistance value will strongly depend on how the conducting regions are interconnected.

FIG. 2. Effective RS methods. (a) A pristine oxide cell. The yellow color represents the high (insulating) resistive area. (b) The case when low resistive (conducting) spots are uniformly generated in the oxide with a volume fraction \( p \). The blue color represents the conducting area. For a very wide range of \( p \), the reduction in resistance is not significant. (c) The case when only the interface of the cell becomes conductive; the cell still has a high resistance. (d) The case when the conducting region has a percolating filamentary shape. This is the most effective method to ensure low cell resistance using only a small volume fraction of the conducting region.

In contrast, Fig. 2(d) shows conducting regions with a filamentary shape. In this case, an interwoven network of conducting regions can be easily formed with a small \( p \). For example, for carbon nanotube composites, a percolating path can be formed only with \( p = 0.016 \).72 In addition, if the filamentary-shaped conducting regions can be located in a limited local region, as shown in the figure, \( p \) can be much smaller. Therefore, this is a very efficient way of inducing a large resistance change. In reality, most RS occurs in this filamentary form, as explained in Sec. III. Throughout this review, we will call these filamentary-shaped conducting regions, generated during RS, conducting filaments (CFs). It should be noted that, in Fig. 2(d), the effective resistance value is determined much more strongly by how the CFs are interconnected to each other than by the local conductivity change.

Because a microscopic theory of RS does not, by itself, take into account how the conducting regions are interconnected, it cannot explain many important physical properties related to RS. For example, most RS phenomena are accompanied by wide fluctuations in various switching parameters, including switching voltages, switching currents, and the resistance of a cell.73–88 These fluctuations in switching parameters should be attributed to the diverse patterns of the conductive paths formed by CFs. Although reducing such fluctuations is the key challenge for commercial memory applications,89 most microscopic theories cannot explain the fluctuations occurring in RS.

At first sight, it is not easy to understand the effects of interconnectivity in CFs even in a qualitative manner, as the patterns usually appear to be random. However, there has been a well-established scientific field relating to percolation theory,90–93 which can quantitatively explain how CFs are interconnected. Specifically, statistical physical concepts, such as scaling, fractals, phase transitions, and critical phenomena, have been well developed to describe the many physical properties of the connectivity of conducting regions in a random inhomogeneous medium.90–97 Therefore, it is natural to apply such well-developed scientific concepts to RS phenomena.

In 2008, Chae et al. proposed a new type of percolation model, called the random circuit breaker (RCB) network model, to explain RS.77 Contrary to the conventional percolation models, which describe either static91,93,94 or irreversible91,93,95–99 phenomena in nature, the RCB network model can describe reversible RS under an electric field. Since this model, there have been numerous developments that have provided deep and quantitative understandings of RS based
on statistical physics approaches, and such activities are increasing. Thus far, none of the earlier reviews has dealt with these important issues systematically. Therefore, it is timely and highly important to review the experimental and theoretical efforts to understand RS from the viewpoint of statistical physics. For this purpose, we provide a review of the recent efforts to understand RS through percolation and scaling theories, both of which are important fields in statistical physics.

In this review, we also deal with the various conversions between switching types observed in one sample. There have been numerous reports of conversion between different switching types (e.g., between unipolar and bipolar switchings\cite{100-107} and between unipolar and threshold switchings\cite{108-112}). However, the understanding of how and why this happens is incomplete. There are two different types of RS with bipolar switching, depending on how the switching between high- and low-resistance states occurs according to polarity. More details on the two types are provided in Sec. II B. There have been several reports that such conversions in bipolar switching can occur even in a single sample.\cite{100-107}

Although the aforementioned conversions may provide crucial insights into the basic mechanisms of RS, there has been minimal effort to explain these in a systematic manner. Here, we systematically review such conversions and propose quantitative explanations based on the mechanism of RS, especially from the viewpoint of statistical physics.

C. Scope and outline of this review

In this review, we wanted to systematically examine RS phenomena, from their microscopic origins to their statistical collective mechanisms, from both qualitative and quantitative perspectives. In Sec. II, we provide an explanation of basic device operations of RS and a classification scheme for RS depending on polarity. In Sec. III, we discuss the microscopic origins of each RS in terms of the above classification. In Sec. IV, we address research efforts to understand RS based on statistical physics approaches such as percolation and scaling theories. In Sec. V, we review research activities related to the conversion between RS types. In Sec. VI, we discuss several technological issues that are closely related those addressed in Secs. IV and V. A more detailed outline of Secs. III–V follows.

In Sec. III, we discuss the microscopic origins of RS based on the role of oxygen vacancies. The general effect of oxygen vacancies on RS and the experimental evidence of oxygen vacancy migration are described in Secs. III A 1 and III A 2, respectively. The microscopic forces driving migration are addressed in Sec. III A 3. In Sec. III B, we discuss unipolar switching, including proposed basic mechanisms for describing processes at the microscopic level and the numerous experimental observations that provide important insight into unipolar switching (Sec. III B 1); the formation process and the nature of CFs (Sec. III B 2); and the rupturing process (Sec. III B 3). In Sec. III C, we discuss microscopic mechanisms for bipolar switching. The causes of polarity-dependent operations are addressed in Sec. III C 1, and the effect of CFs on bipolar switching is discussed in Sec. III C 2. In Sec. III C 3, we discuss the microscopic mechanisms of the set and reset processes in bipolar switching.

In Sec. IV, we present an overview of the statistical physics approaches to unipolar switching. In Sec. IV A, we discuss the limitation of the microscopic studies described in Sec. III and then describe why statistical physics approaches are important for understanding the fluctuations of switching parameters, such as switching current and voltages, observed in RS phenomena. We explain percolation models for unipolar switching, including the RCB network model, in Sec. IV B. In Sec. IV C, we survey the scaling behaviors originating in the fluctuations in switching parameters and review scaling theory to understand the scaling behaviors.

In Sec. V, we review research activities relating to switching-type conversion problems. In Sec. V A, conversion between unipolar and threshold switchings is discussed. In Sec. V B, conversion between unipolar and bipolar switchings is reviewed, and the percolation model approach to bipolar switching is explained. In Sec. V C, we provide a review of switching-polarity conversion in bipolar switching.

In Sec. VI, we review complex RRAM technological issues. The potential applications of RRAM have led to renewed interest in RS, and there have been a number of recent important technological developments, which we preliminarily describe. Then, of the many RRAM technological issues, we review those closely related to the aforementioned statistical physics perspective. We focus particularly on switching voltage distribution in unipolar switching and how to improve the fluctuation of the switching parameters. Finally, in Sec. VII, we discuss the outlook for RS applications and conclude the paper.

II. RESISTIVE SWITCHING AND ITS CLASSIFICATION

A. Basic operation of electric field-induced resistive switching

Typical devices for RS applications have a capacitor-like two-terminal configuration, as shown in Fig. 3(a), where an oxide thin film is sandwiched between top and bottom electrodes. This capacitor-like configuration is convenient, as a large electric field (typically \(\sim 10\%\) of the dielectric strength,\cite{113} which in many solids is of the order of \(10–100\) MV/m) can be applied easily. In addition, this configuration is relatively easy to fabricate, which is one of the advantages for its use in practical RRAM devices. However, it is somewhat difficult to investigate internal structural changes in the film during RS; for example, it is quite difficult to directly observe changes in the percolating CFs formed inside the film under external bias. For this reason, some researchers have used an in-plane configuration, as shown in Fig. 3(b), rather than a capacitor-like configuration to better observe the changes that the material undergoes during RS.\cite{37,44,114-116}

Figure 3(c) shows the basic operation of an RS device. An oxide insulator should have a very large resistance because the energy gap between the conduction and valence bands (typically, \(1–4\) eV) is much larger than the thermal energy at room temperature (\(\sim 25\) meV). Therefore, in the pristine state, most oxide films have a very large resistance, as long as the chemical composition is close to the...
stoichiometric value. However, under a large external voltage, $V_{\text{ext}}$, the device enters a low-resistance state (LRS). This process is called “electroforming” or “forming,” indicated by the green curve in the figure, and the required voltage is termed the “forming voltage,” $V_{\text{forming}}$. During the forming process, it is important to limit the flowing current; otherwise, the film will experience a complete dielectric breakdown. To prevent such irreversible damage, there should be a maximum value of flowing current, which is termed the compliance current.117 Note that the forming process is a prerequisite for most oxides to exhibit RS. We discuss the origin of the forming process in Sec. III B 2.

Following the forming process, the resistance of the device can be switched back and forth between the LRS and a high-resistance state (HRS) by controlling $V_{\text{ext}}$. In the LRS, as $V_{\text{ext}}$ increases, a sudden increase in resistance occurs, as shown by the red curve in Fig. 3(c). This is called the “reset” process, and the corresponding voltage is called the “reset voltage,” $V_{\text{reset}}$. The HRS is metastable and typically exhibits non-metallic $I$–$V$ characteristics. When we increase $V_{\text{ext}}$ from zero once more, the device can be switched back to the LRS when $V_{\text{ext}}$ reaches the set voltage, $V_{\text{set}}$, as shown by the blue curve in Fig. 3(c). This process is called the “set” process. Note that the resistance changes during the reset and set processes can be either abrupt or gradual. Similar to the forming process, during the set operation, a setting of the compliance current is usually required to avoid a complete dielectric breakdown. The resulting LRS after the set operation is also stable, and the set and reset switching processes can be repeated many times. Although the resistance values of LRS and HRS are always accompanied by sizable fluctuations, the magnitude of the change in resistance remains within well-defined limits. Therefore, the two metastable states (i.e., the LRS and HRS) can be used as binary states for nonvolatile memory applications.

In practical operations, the details of the $I$–$V$ curves are important. Note that the memory operation depends on the relative polarity of $V_{\text{reset}}$ and $V_{\text{set}}$. If they have the same polarity, we can use only positive or negative $V_{\text{ext}}$ for memory operations, called “unipolar.” On the other hand, if they have different polarities, we should use both positive and negative $V_{\text{ext}}$, called “bipolar.” The required polarities of $V_{\text{forming}}$, $V_{\text{reset}}$, and $V_{\text{set}}$ have been observed to differ from one material to another. Therefore, it is convenient to classify RS phenomena as either unipolar or bipolar (depending on the required polarity operations), as explained in Sec. II B.

B. Classification of resistive switching phenomena

A classification scheme for RS is useful for obtaining a thorough understanding of the physics of RS phenomena. In this review, we classify RS into three groups: unipolar, bipolar, and threshold switchings. Unipolar and bipolar switchings exhibit at least two stable states without an applied bias, and these are therefore suitable for non-volatile memory applications. Threshold switching can also have multiple states in response to an applied bias; however, there is only one stable state when the external voltage is zero. The memory function is therefore volatile. Table I lists materials that exhibit unipolar, bipolar, and threshold switchings. As the table shows, unipolar switchings are usually observed in binary transition metal oxides, such as NiO$^{24–26,30,73–75,108,109,118–128}$ TiO$^{27,78,129–134}$ and Nb$_2$O$_5$.135 On the other hand, bipolar switching has been reported in ternary$^{19,136–149}$ as well as some binary oxides.144,171–174 Interestingly, only bipolar switching has been reported in single crystal oxides.144,171–174 This table also includes the materials that show both unipolar and bipolar switchings in one cell. The coexistence of unipolar and bipolar switchings has been
observed in some oxides, such as TiO₂, SrTiO₃, NiO, and so on. Threshold switching has been observed in binary oxides. In Secs. II B 1–II B 3, we describe the basic features of I–V curves for each type of RS phenomena.

1. Unipolar switching

Figure 4(a) shows a typical I–V relation for unipolar switching after the oxide sample experienced the forming process. Consider, for example, that the sample is in the LRS and a positive external bias \( V_{\text{ext}} \) (>0) is applied to it. When the voltage \( V_{\text{reset}} \) is reached, the resistance abruptly increases and the device enters the HRS. This is the reset process during which the compliance current limit is usually not required, because the high resistance value of the HRS will self-limit the current flow. The HRS is also metastable and remains for a long period of time with no applied bias. When a positive \( V_{\text{ext}} \) is applied to the sample in the HRS, an abrupt reduction in resistance occurs at \( V_{\text{set}} \), and the sample goes into the LRS again. This is the set process during which the compliance current limits should be set. Typically, \( V_{\text{set}} \) is larger than \( V_{\text{reset}} \); an exception was observed in a SiOₓ device. Note, that for unipolar switching, the I–V curve is symmetric about \( V_{\text{ext}} \). Namely, when a negative bias is applied, similar reset and set processes also occur. Therefore, external voltage pulses with only one polarity are needed to change the resistance state. To prevent a permanent dielectric breakdown, a compliance current is needed. In all observed experiments, the resistance change during set and reset was discontinuous. With a few exceptions, the reset voltage was generally smaller than the set voltage. (b) Schematic of a typical figure-of-eight (F8) bipolar I–V curve, which is asymmetric with respect to the polarity of an external voltage. Therefore, set and reset both occur with different polarities. In some cases, a compliance current is not needed and the resistance change during set and reset is not sudden, but gradual. (c) Schematic of a typical counter-figure-of-eight (cF8) bipolar I–V curve: the set and reset occur at a different polarity compared with the F8 case. (d) Schematic of a typical (unipolar) threshold I–V curve, which is symmetric with respect to the polarity of an external voltage. The LRS is maintained only when a bias is applied.

<table>
<thead>
<tr>
<th>Type</th>
<th>Proposed mechanism</th>
<th>Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unipolar switching</td>
<td>Filament formation/rupture</td>
<td>Cu-doped ZrO₂, CuOₓ, FeOₓ, MnOₓ, NiO, ZnO, CoOₓ, NbOₓ, SrTiO₃, NiO, SrTiO₃, Cu-doped SiO₂</td>
</tr>
<tr>
<td>Bipolar switching</td>
<td>Charge trap/detrap</td>
<td>WOₓ, Al₂O₃, SrTi₁₋ₓNbₓO₃, single crystal, Sm₂O₃, CaₓMnO₃, Pr₀.₇C₀.₃MnO₃, Cr-doped SiZrO₃, Pt-dispersed SiO₂</td>
</tr>
<tr>
<td>Oxygen vacancy migration</td>
<td></td>
<td>Cr₂O₃, graphene oxide, TiO₂, WOₓ, BiFeO₃, Fe-doped SrTiO₃, Fe-doped CuMnO₃, Ta₂O₅, Ta₂O₅/C₀ₓ, Ta₂O₅/TaOₓ/TiO₂</td>
</tr>
<tr>
<td>Threshold switching</td>
<td></td>
<td>NiO, Coexistence of unipolar and bipolar switchings</td>
</tr>
<tr>
<td>Coexistence of unipolar and bipolar switchings</td>
<td></td>
<td>Ta₂O₅/TaOₓ, TiO₂, Ba₀.₇Sr₀.₃TiO₃, Cu-doped SiO₂, NiO, HfO₂</td>
</tr>
<tr>
<td>Coexistence of unipolar and threshold switchings</td>
<td></td>
<td>NiO, SrTiO₃, (Ba,Sr)TiO₃, HfO₂</td>
</tr>
</tbody>
</table>

FIG. 4. (a) Schematic of a typical unipolar (or nonpolar) I–V curve, which is symmetric with respect to the polarity of an external voltage. Thus, set and reset can occur with a single polarity. To prevent a permanent dielectric breakdown, a compliance current is needed. In all observed experiments, the resistance change during set and reset was discontinuous. With a few exceptions, the reset voltage was generally smaller than the set voltage. (b) Schematic of a typical figure-of-eight (F8) bipolar I–V curve, which is asymmetric with respect to the polarity of an external voltage. Therefore, set and reset both occur with different polarities. In some cases, a compliance current is not needed and the resistance change during set and reset is not sudden, but gradual. (c) Schematic of a typical counter-figure-of-eight (cF8) bipolar I–V curve: the set and reset occur at a different polarity compared with the F8 case. (d) Schematic of a typical (unipolar) threshold I–V curve, which is symmetric with respect to the polarity of an external voltage. The LRS is maintained only when a bias is applied.
sufficient in real device operations. This is why such RS is called “unipolar.”

2. Bipolar switching

With bipolar switching, both polarities are required. One form of a typical I–V curve for bipolar switching is shown in Fig. 4(b). A negative bias is required for the reset process, which takes the device from the LRS into the HRS. Then, a positive bias is required for the set process, which takes the device from the HRS into the LRS. Because both polarities of \( V_{\text{ext}} \) are required, this type of RS is called bipolar switching.

Note that there are two possible types of bipolar switching. Figure 4(b) shows an I–V curve in which the reset (set) process occurs with a negative (positive) bias. Such a scheme is referred to as “figure-of-eight (F8)” bipolar switching, as the corresponding I–V curve follows the pattern of writing the number “8.” Some bipolar switching systems employ a positive bias for the reset process and a negative bias for the set process, as schematically shown in Fig. 4(c). The corresponding I–V curve is referred to as a “counter-figure-of-eight (cF8)” bipolar switching curve. Table II lists materials that exhibit cF8 and F8 switching schemes. Both cF8 type and F8 type bipolar switching occur in many oxide materials. Some materials, such as \( \text{WO}_3 \), \( \text{Pr}_0.7\text{Ca}_{0.3}\text{MnO}_3 \) thin films, and \( \text{Nb-doped SrTiO}_3 \) single crystals show both types of bipolar switching. The existence of these two types of bipolar switching is confusing, but it turns out to be closely related to their microscopic origin. This issue will be addressed in detail in Sec. VI C.

For unipolar switching, the forming process is always required. However, the requirement for the forming process in bipolar switching varies depending on the materials system. Such a variation has caused much confusion. With bipolar switching, two types of conducting region geometry have been reported (i.e., Figs. 2(c) and 2(d)). In some papers, these were called homogeneous and inhomogeneous bipolar switching, respectively. However, in statistical physics, the term “inhomogeneous” is usually used to describe how a locally averaged macroscopic property varies in space. When the macroscopic conductivity varies along the vertical direction, the geometry of Fig. 2(c) can be inhomogeneous. Therefore, we will use different terminology in this review. For Fig. 2(c), we will say, that switching occurs laterally uniformly along the interface. When bipolar switching occurs laterally nonuniformly along the interface, it is usually necessary to form local CFs inside the material, which is similar to unipolar switching. The effect of CFs will be discussed in detail in Secs. III C 2 and III C 4. However, for a very thin film, typically less than tens of nanometers, laterally uniform bipolar switching can occur, and the forming process may not be required. When the forming process is required, the compliance current limit typically needs to be set during the operation of the device.

3. Threshold switching

Threshold switching describes RS phenomena for which there is only one stable state with no external bias. Figure 4(d) shows a typical I–V curve for threshold switching. At \( V_{\text{set}} \), the device changes from the HRS to the LRS. The LRS is stable over only a certain range of applied biases, and when the applied bias falls below this range, the device reverts to the HRS.

Although threshold switching occurs much less frequently than unipolar or bipolar switching, it can be a useful phenomenon, scientifically as well as technologically. For example, the I–V curve is symmetrical with polarity: in this regard, threshold switching is similar to unipolar switching. In Sec. V A, we address how the conversion between unipolar memory and threshold switchings can occur, which can provide insights into the role of Joule heating in both types of RS. Due to the instability of some regions of CFs in LRS, threshold switching devices are not suitable for non-volatile memory applications. However, many potential technological applications using threshold switching have been proposed, including electrical switches, smart windows, terahertz nanoantennas, and memory metamaterials. In addition, it was recently proposed that threshold switching be used as a selector to solve a sneak path problem occurring in a RRAM crossbar array.

### TABLE II. Type classification of bipolar switching and corresponding materials.

<table>
<thead>
<tr>
<th>Type</th>
<th>Proposed mechanism</th>
<th>Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>F8</td>
<td>Charge trap/detrap Oxygen vacancy migration</td>
<td>( \text{WO}_3 ), ( \text{SrTiO}_3 ), ( \text{NbO}<em>3 ), ( \text{SrTiO}<em>3 ) single crystal, ( \text{Sm}</em>{0.7}\text{Ca}</em>{0.3}\text{MnO}_3 ), ( \text{Pt}<em>0.7\text{Ca}</em>{0.3}\text{MnO}_3 ) thin film, ( \text{BiFeO}_3 ), ( \text{Fe-doped SrTiO}<em>3 ) thin film, ( \text{Pr}<em>0.7\text{Ca}</em>{0.3}\text{MnO}<em>3 ) thin film, ( \text{Ta}</em>{0.7}\text{Nb}</em>{0.3}\text{O}_3 ), ( \text{TaO}_2 ), ( \text{HFO}_2 ), ( \text{Al}_2\text{O}_3 ), Cr-doped ( \text{SrTiO}_3 ) single crystal, ( \text{Al}_2\text{O}_3 ), Cr-doped ( \text{SrZrO}_3 ), Pt-dispersed ( \text{SiO}_2 ) thin film, ( \text{TiO}<em>2 ), ( \text{TiO}<em>2 / \text{SiO}<em>2 ) stacked film, ( \text{TiO}<em>2 / \text{TiO}<em>2 ) stacked film, ( \text{Cr}-\text{doped SrTiO}<em>3 ) single crystal, ( \text{Pr}</em>{0.7}\text{Ca}</em>{0.3}\text{MnO}<em>3 ) thin film, ( \text{Ta}</em>{0.7}\text{O}</em>{2-x} / \text{Ta}</em>{0.7}\text{O}</em>{2-x} ) stacked film, ( \text{SrTiO}<em>3 ) single crystal, ( \text{Al}</em>{2}\text{O}</em>{3} ) single crystal</td>
</tr>
<tr>
<td>cF8</td>
<td>Charge trap/detrap Oxygen vacancy migration</td>
<td>( \text{WO}_3 ), ( \text{SrTiO}_3 ), ( \text{NbO}<em>3 ), ( \text{SrTiO}<em>3 ) single crystal, ( \text{Sm}</em>{0.7}\text{Ca}</em>{0.3}\text{MnO}_3 ), ( \text{Pt}<em>0.7\text{Ca}</em>{0.3}\text{MnO}_3 ) thin film, ( \text{BiFeO}_3 ), ( \text{Fe-doped SrTiO}<em>3 ) thin film, ( \text{Pr}<em>0.7\text{Ca}</em>{0.3}\text{MnO}<em>3 ) thin film, ( \text{Ta}</em>{0.7}\text{Nb}</em>{0.3}\text{O}_3 ), ( \text{TaO}_2 ), ( \text{HFO}_2 ), ( \text{Al}_2\text{O}_3 ), Cr-doped ( \text{SrTiO}_3 ) single crystal, ( \text{Al}_2\text{O}_3 ), Cr-doped ( \text{SrZrO}_3 ), Pt-dispersed ( \text{SiO}_2 ) thin film, ( \text{TiO}<em>2 ), ( \text{TiO}<em>2 / \text{SiO}<em>2 ) stacked film, ( \text{TiO}<em>2 / \text{TiO}<em>2 ) stacked film, ( \text{Cr}-\text{doped SrTiO}<em>3 ) single crystal, ( \text{Pr}</em>{0.7}\text{Ca}</em>{0.3}\text{MnO}<em>3 ) thin film, ( \text{Ta}</em>{0.7}\text{O}</em>{2-x} / \text{Ta}</em>{0.7}\text{O}</em>{2-x} ) stacked film, ( \text{SrTiO}<em>3 ) single crystal, ( \text{Al}</em>{2}\text{O}</em>{3} ) single crystal</td>
</tr>
</tbody>
</table>
A Schottky barrier is explained in Sec. III C. The detailed RS mechanism due to the modulation of the oxide layer’s work functions between the metal electrode and the oxide layer is shown. Depending on the difference of the electrodes and the oxide layer, the Schottky barrier can be modulated by the neutralization of the oxygen vacancies due to the trapping of electrons. The position of the oxygen vacancies at each resistance state was investigated by using electron energy-loss spectra (EELS). Figure 5(a) shows the EELS mapping images of the O–K (O 1s → 2p, 532 eV) edge in the LRS and HRS. When the cell was in the LRS (upper panel), most of the Nb: SrTiO₃ region had uniformly bright spots, representing an almost stoichiometric oxygen occupancy. However, when the cell was in the HRS (lower panel), darker spots were observed at a depth of ~10 nm beneath the electrode. Since the contrast in the EELS image originates primarily from the intensity contrast, the darker spots indicated that the stoichiometry of the region had deviated from the ideal stoichiometry. Indeed, the oxygen vacancy concentration just under the Pt top electrode became considerably larger in the HRS than in the LRS, which was also confirmed from measurements of the oxygen intensity using secondary ion mass spectrometry. These experiments clearly showed that the oxygen vacancies migrate inside the oxide cell under an electric field and that such movement can induce resistance changes.

Other studies have been carried out to directly detect and/or visualize the migration of oxygen vacancies inside an oxide sample. One such experiment was performed using spectroscopic tools, which can provide the local chemical stoichiometric information directly. Janousch et al. used micro X-ray fluorescence and X-ray absorption near-edge spectroscopy on a Cr-doped SrTiO₃ single crystal cell and found that an oxygen vacancy-rich region had formed near the electrodes after a forming process. Another interesting experiment was performed using electrocoloration, also known as electrochromism. Under an electric field, the color of some materials (i.e., SrTiO₃ single crystal [198] and BaTiO₃ single crystal [199]) is known to change reversibly due to electrochemical redox reactions. Waser et al. used an Fe-doped SrTiO₃ single crystal [200] which has different optical properties depending on the valence state of Fe. They were able to visualize the oxygen vacancy concentration using a color gradient. When an external electric field was applied to the single crystal for a few minutes, the color near the cathode became brighter because positively charged oxygen vacancies migrated toward the cathode; thus, the concentration of Fe⁺ decreased there.

1. How do oxygen vacancies play a role in resistive switching?

In general, oxygen vacancies will affect the resistance and contribute toward the induction of RS phenomena in one of the following mechanisms: (1) Oxygen vacancies tend to cluster and generally form filamentary shapes under an electric field. When such clusters are formed, the resistance of the local region becomes much lower than that of the surrounding oxide matrix, and the LRS and HRS will therefore be determined by the creation and rupture of the CFs, respectively, which percolate through the sample. The relevant models are explained in Sec. III B. (2) Oxygen vacancies can control the Schottky barrier characteristics. In the capacitor sample geometry, as shown in Fig. 3(a), an interface is formed between the electrodes and the oxide layer. Depending on the difference of the work functions between the metal electrode and the oxide, a Schottky barrier can be formed. Under an external electric field, the distribution and density of the oxygen vacancies can vary, which affects the height and width of the Schottky barrier. The detailed RS mechanism due to the modulation of the Schottky barrier is explained in Sec. III C. (3) The claim has been made that oxygen vacancies can also form trap sites for electrons inside the Schottky barrier region. In this case, the Schottky barrier can be modulated by the neutralization of the oxygen vacancies due to the trapping of electrons, which also leads to RS phenomena. The models based on electron traps by the oxygen vacancies will be covered in Sec. III C.3.

2. Evidence of oxygen vacancy migration

a. Microscopic observations. Ions can migrate in a solid. Most metallic ions have a higher atomic number Z than oxygen ions, and thus generally have a much larger interaction cross section to incoming electrons and X-rays. In the case of metallic ions, their movement inside a material can be relatively easily observed using direct microscopic imaging techniques such as transmission electron microscopy (TEM) or scanning electron microscopy (SEM). However, the observation of oxygen vacancy migration inside a bulk oxide is much more difficult than that of metallic ions due to its small Z. Muller et al. reported that oxygen vacancy defects below 8% in SrTiO₃ thin films could not be observed using high-angle annular dark-field images because their signals were proportional to Z³. Below this threshold, typical oxide thin films, the concentration of oxygen vacancies is too small to be detected using direct imaging techniques such as TEM. However, at the oxide surface or interface, oxygen vacancy movement will be blocked and they will accumulate. Therefore, the concentration of oxygen vacancies could be enhanced significantly. For RS, several studies have reported the movement of oxygen vacancies in an oxide surface.

A recent systematic study of the migration of oxygen vacancies and how such movement affects the resistance of a material has been undertaken in an Nb-doped SrTiO₃ single crystal cell. In that study, a Pt/Nb: SrTiO₃/Ti cell was used, which exhibited bipolar switching with an F8 polarity. The position of the oxygen vacancies at each resistance state was investigated by using electron energy-loss spectra (EELS). Figure 5(a) shows the EELS mapping images of the O–K (O 1s → 2p, 532 eV) edge in the LRS and HRS. When the cell was in the LRS (upper panel), most of the Nb: SrTiO₃ region had uniformly bright spots, representing an almost stoichiometric oxygen occupancy. However, when the cell was in the HRS (lower panel), darker spots were observed at a depth of ~10 nm beneath the electrode. Since the contrast in the EELS image originates primarily from the intensity contrast, the darker spots indicated that the stoichiometry of the region had deviated from the ideal stoichiometry. Indeed, the oxygen vacancy concentration just under the Pt top electrode became considerably larger in the HRS than in the LRS, which was also confirmed from measurements of the oxygen intensity using secondary ion mass spectrometry. These experiments clearly showed that the oxygen vacancies migrate inside the oxide cell under an electric field and that such movement can induce resistance changes (see also Sec. V C.3 for an explanation of the relationship between resistance and the oxygen vacancy concentration).

Other studies have been carried out to directly detect and/or visualize the migration of oxygen vacancies inside an oxide sample. One such experiment was performed using spectroscopic tools, which can provide the local chemical stoichiometric information directly. Janousch et al. used micro X-ray fluorescence and X-ray absorption near-edge spectroscopy on a Cr-doped SrTiO₃ single crystal cell and found that an oxygen vacancy-rich region had formed near the electrodes after a forming process, as shown in Fig. 5(b).

Another interesting experiment was performed using electrocoloration, also known as electrochromism. Under an electric field, the color of some materials (i.e., SrTiO₃ single crystal [198] and BaTiO₃ single crystal [199]) is known to change reversibly due to electrochemical redox reactions. Waser et al. used an Fe-doped SrTiO₃ single crystal [200] which has different optical properties depending on the valence state of Fe. They were able to visualize the oxygen vacancy concentration using a color gradient. When an external electric field was applied to the single crystal for a few minutes, the color near the cathode became brighter because positively charged oxygen vacancies migrated toward the cathode; thus, the concentration of Fe⁺ decreased there.
whose CFs are formed by metal ions (Ag\textsuperscript{190,203,204} or Cu\textsuperscript{205–208}) dissolved from electrodes. Before discussing the oxygen vacancy migration, we first review metal ion migration in the electrochemical metallization cell. The electrochemical metallization cells make use of electrochemical aggregation and dissolution of metal ions for RS operations.\textsuperscript{38,44,190,201–205,207–210} The cells typically have an electrode/insulator/electrode structure, in which one electrode is an electrochemically active metal and the other is an electrochemically inert metal. The insulator is a thin film of a solid electrolyte. When an electric field is applied to the electrolyte and then aggregate near the inert electrode. Such migration and aggregation of active metal ions can be directly observed by using TEM and SEM.\textsuperscript{189,190}

Figure 6(a) shows the electric-field dependence of the wait time measured from Ag/SiO\textsubscript{2}/Pt electrochemical metallization cells. Here, the wait time \( t_W \) is the duration between the voltage being applied to the cell and a sudden RS event occurring. The wait time decreases exponentially as the electric field increases. This switching-time dependence can be understood from the ionic motion inside an electrolyte.\textsuperscript{190} Although some debate exists as to whether migration or aggregation is the rate-limiting process,\textsuperscript{207} the exponential decrease in the wait time can be explained in terms of ionic transport. Consider, for example, a periodic potential that is inclined when an electric field \( E \) is applied to the device, as shown in Fig. 7(a). If we assume that ionic transport is the rate-limiting process, the wait time \( t_W \) can be expressed as

\[
 t_W \propto \frac{1}{v} e^{\beta(E_a - qaE)}, \tag{1}
\]

where \( v \) is the attempt frequency; \( \beta \) is the inverse temperature, i.e., \( 1/k_B T \), where \( k_B \) is Boltzmann's constant and \( T \) is the temperature; \( E_a \) is the activation energy; \( a \) is the lattice constant; and \( q \) is the charge of the ion. The exponential decrease in \( t_W \) as a function of the electric field [i.e., following Equation (1)] has been widely accepted as evidence of ionic transport.

FIG. 5. (a) The upper and lower panels show electron energy-loss spectra (EELS) O-\textit{K} (O 1\textit{s} \rightarrow 2\textit{p}, 532 eV) edge mapping images of the LRS and the HRS, respectively. Reproduced with permission from Lee et al., APL Mater. 2, 066103 (2014). Copyright 2014 AIP Publishing. (b) An X-ray fluorescence map of Cr in Cr-doped SrTiO\textsubscript{3} resistance change memory. The red color represents oxygen vacancies in the Cr octahedral. Reproduced with permission from Janousch et al., Adv. Mater. 19, 2232 (2007). Copyright 2007 John Wiley & Sons, Inc.

FIG. 6. The switching dynamics during RS processes. (a) The wait-time dependence on an external electric field as observed in an Ag/SiO\textsubscript{2}/Pt device. In this case, the wait time is the duration of time before a sharp resistance-switching event occurs. Reprinted with permission from Yang et al., Nat. Commun. 3, 732 (2012). Copyright 2012 Macmillan Publishers. (b) The wait-time dependence on an applied pulse amplitude observed in Pt/SrTiO\textsubscript{x}/Pt cells. Reprinted with permission from Lee et al., Appl. Phys. Lett. 98, 053503 (2011). Copyright 2011 AIP Publishing.
A similar switching-time dependence was observed in oxide RS cells. Figure 6(b) shows the dependence of the waiting time on the electric field in Pt/SrTiO$_3$/Pt cells. In this case, the waiting time is the forming time, which is defined as the time from the application of an external voltage pulse to the occurrence of an abrupt increase in the current. The figure shows that the waiting time decreases exponentially with increasing pulse amplitude, supporting the theory that ionic migration is involved in RS phenomena in oxides. The oxygen-vacancy motion under no external bias can be described as random hopping from one site to a neighboring site, with an energy barrier of a few electron volts. This energetic barrier for hopping declines when an external electric field is applied. At this point, the oxygen vacancies migrate in one direction by hopping, as shown in Fig. 7(a), which is governed by the Arrhenius equation (1). Therefore, the observed switching-time dependence in oxide RS cells strongly supports oxygen vacancy migration during RS.

The migration of oxygen vacancies under an electric field can be reasonably described by the thermally activated process in Equation (1). However, more than one exponential regime exists, as shown in Fig. 6(b), which strongly suggests the presence of other competing microscopic origins for oxygen vacancy motion. A full understanding of such an oxygen vacancy mechanism is important in understanding RS. Therefore, in Sec. III A 3, we review three important origins of the microscopic forces that determine the dynamics of oxygen vacancy migration.

3. Microscopic origins for oxygen vacancy migration

a. Electric field. The primary source of the microscopic forces inducing oxygen vacancy migration in RS is the electric field. Since oxygen vacancies are positively charged and oxygen ions are negatively charged, oxygen vacancies will move toward the cathode and oxygen ions will move toward the anode. The resulting ionic motion is therefore parallel to the applied electric field. When the matrix oxide is crystalline, the migration of oxygen vacancies can be thought of as a hopping process in a periodic potential, as shown in Fig. 7(a), called the “washboard potential.” When an electric field $E$ is applied, the hopping rate $r_{\text{hop}}$ is proportional to the inverse of $t_w$ [see Equation (1)],

$$r_{\text{hop}} = ve^{-\beta(E_a - qAE)},$$

The attempt rate is typically $v \sim 10^{13}$ Hz and the typical activation energy is $E_a \sim 1.0$ eV.

Note that at room temperature, the effect of the electric field on the process of hopping is not sufficiently large to explain the actual hopping occurring inside RS materials. For example, when $q = 2$, $E \sim 0.1$ V/nm and $a \sim 0.5$ nm, and at room temperature, $\beta \sim 40$ eV$^{-1}$; the effect of the electric field is quite small, giving $r_{\text{hop}} \sim 10^{-4}$ s$^{-1}$, or approximately one hop per hour. A small value of $r_{\text{hop}}$ will result in a long retention time and very slow switching during the write/erase processes. The latter is not compatible with next-generation memory device requirements and so this issue has been termed the “voltage–time dilemma.” However, RS phenomena actually occur on a much faster timescale; even (sub)nanosecond switching has been reported. It follows that the electric field cannot be the only force that can induce ion migration, and some other acceleration mechanism is required for fast switching.

One such acceleration mechanism is local Joule heating. Under the external bias, a current flows in the conducting channel, so the temperature is able to rise locally near to the channel. Some studies have reported that Joule heating can raise the local temperature by several hundred Kelvins near the CFs. At such temperatures, the hopping rate becomes accelerated, as in Eq. (2), and can reach $\sim 10^7$ s$^{-1}$. This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 147.46.19.34 On: Tue, 13 Oct 2015 06:35:15
This $r_{\text{hop}}$ value is comparable to the observed nanosecond switching speed of RS. In addition to this acceleration mechanism, the temperature gradient generated by Joule heating can provide another way to affect the motion of ions, as described in Sec. III A 3 b.

b. Soret force due to Joule heating. Without an electric field, the temperature gradient $\nabla T$ can result in anisotropic hopping, giving rise to a net motion of ions. As shown in Fig. 7(b), an oxygen ion can be expected to move from a hotter to a cooler region because the thermal energy of the particle in the hotter region is greater than that in the cooler region. Since the average velocity of the oxygen ion is higher in the hotter region, oxygen vacancies are expected to move from cooler to hotter regions. Such a motion resulting from $\nabla T$ can be represented by a microscopic force, termed a Soret force.

The Soret force can act as a microscopic force that attracts oxygen vacancies, giving rise to regions that are oxygen vacancy-rich during the forming and set processes. Some researchers have claimed that the forming and set processes consist of two steps. The first step is responsible for the formation of electrical current filaments, shown by the yellow curves in Fig. 7(d). Note that this electrical filament is a temporary pathway of electrons without atomic rearrangement, so it is different from the usual nonvolatile CF. The argument has been made that the electrical filaments are formed due to “soft electrical breakdown”, thus, they disappear when an electrical stimulus is disrupted. These electrical filaments give rise to significant Joule heating, which in turn creates a large $\nabla T$. Since the average direction of the electrical filament should be parallel to the applied electric field, $\nabla T$ will be almost perpendicular to the electric field. It follows that the Soret force acts effectively in a perpendicular direction to the electric field. This force will attract oxygen vacancies to the electrical filament and form an oxygen vacancy-rich region, which makes nonvolatile percolating CFs. Since $\nabla T$ does not depend on the current direction, the Soret force is therefore polarity-independent and should thus play an important role in unipolar switching phenomena.

c. Restoring or Fick force due to Joule heating. Joule heating can provide another microscopic force, which is the opposite of the Soret force, and can restore the oxygen vacancy-rich region to its original configuration. When the oxygen vacancy concentration $n$ is highly non-uniformly distributed, nature will try to make $n$ uniform to minimize the free energy (i.e., to maximize the entropy term). Thus, oxygen vacancies tend to move from regions of higher concentration to regions of lower concentration owing to diffusion, as shown in Fig. 7(c). Such motion due to $\nabla n$ can be effectively described in terms of a microscopy force, called the Fick force, which becomes important in regions of high temperature because the hopping rate for ions depends exponentially on the temperature, as described in Sec. III A 2 b.

Note that the Fick force plays an important role during the reset process. When the oxide is in the original atomic configuration (i.e., the pristine state), $n$ oxygen vacancies will be roughly uniform throughout the material. Therefore, no diffusive flow of ions can occur due to the negligible $\nabla n$ of the ion concentration. On the other hand, when the oxide is LRS after forming or set processes, percolating CFs appear close to where many oxygen vacancy clusters are formed. The concentration in those regions then becomes larger than in the surrounding oxide matrix and $\nabla n$ develops. With Joule heating, the Fick force dominates and oxygen vacancies tend to move from regions that are rich in vacancies into the surrounding oxide. Thus, the LRS changes to become the HRS.

In Secs. III A 3 a–III A 3 c, we described three microscopic forces that can give rise to ion migration: the electric field, the Soret force (resulting from Joule heating), and the Fick force. In general, these three forces may be present simultaneously, as shown in Fig. 7(d). Note that the current flow induced by the electric field gives rise to Joule heating simultaneously in RS phenomena. Competition among these three forces will then determine the resulting motion of the ions inside the oxide. However, a detailed microscopic description of the motion of these ions for a particular type of switching process still remains to be studied. Experiments to visualize the atomic structure in real time during switching using, for example, in situ TEM, will provide valuable information. Theoretical studies based on molecular dynamics are also highly valuable to better understand the microscopic details regarding the motion of ions during switching.

B. Unipolar switching

1. What happens in an oxide during unipolar switching?

a. Microscopic studies. Many microscopy experiments have been carried out to gain insights into the RS mechanisms involved in unipolar switching. Figure 8 shows a conductive atomic force microscopy (C-AFM) experiment on a TiO$_2$ film, while Fig. 8(a) presents a schematic diagram of the C-AFM measurements in which the oxide film is located on a Pt bottom electrode, and the C-AFM tip is used as the top electrode. Figure 8(b) shows $I$–$V$ curves measured at a single point on the oxide surface. Application of $V_{\text{ext}}$ to the C-AFM tip induces unipolar switching, and thus, forming, reset, and set processes can be achieved. The C-AFM experiment provides information on how the conductive percolating path will touch the top surface with a high lateral spatial resolution. Namely, it shows the top cross section of the local current pathways formed inside the oxide film. Figure 8(c) shows current maps of the oxide surface for the LRS (i.e., after the forming process), and Fig. 8(d) displays maps for the HRS (i.e., after the reset process). In the LRS, several highly conducting regions were observed that were $3$–$10$ nm across, while in the HRS, these conducting regions disappeared. The experiments show that local conductive pathways between the bottom electrode and the top surface should be present in the LRS, but should be disrupted in the HRS.
In a similar experiment, liquid Hg was used as a removable top electrode for a NiO film. Following switching of the NiO film to the LRS or HRS, the top liquid electrode was able to be removed, and current maps were then measured using C-AFM. This experiment also supported the existence of local current pathways in the LRS, which were disrupted by the reset process. However, such C-AFM studies using the capacitor geometry cannot describe the detailed changes occurring inside the bulk.

Using the in-plane geometry described in Fig. 3(b), more physical insights can be obtained on the percolating conductive pathways. Figure 9 shows SEM images of a planar Ni/CuO/Ni cell. As shown in Fig. 9(a), after the forming process a bridge-like structure was observed in the CuO, which corresponds to the LRS. The structure was then cut using a focused ion beam, as displayed in Fig. 9(b). The resistance of the cell increased from 3.3 kΩ to 196 kΩ upon entering into the HRS, which indicates that the bridge between the conducting Ni electrodes was a conductive filament channel. Subsequently, an electric field was applied again to the cell. As shown in Fig. 9(c), another bridge structure formed and the resistance of the cell dropped, returning the cell to the LRS. This work clearly demonstrated that unipolar switching occurs by the formation and rupture of percolating CFs inside a sample.

Recent in situ TEM measurements have enabled observations of the filament formation process in real time. Figure 10 shows the in situ TEM images, while generating and disrupting a percolating CF in a ZnO film during RS operations. The upper panels of Fig. 10 reveal sequential TEM images of the growth of a CF during the forming process. The boundary of the filament is indicated by a white dashed line. Note that an intermediate dendritic structure finally percolated through the oxide. The lower panels show sequential TEM images of the reset process during which the CF was partially ruptured—consistent with the sudden disappearance of the conductive spikes observed in the C-AFM experiments following switching to the HRS.

Based on the observations of these C-AFM, SEM, and in situ TEM experiments, unipolar switching appears to be due to the formation and rupture of the CFs, as shown schematically in Fig. 11. As revealed in Fig. 11(a), no CF is present inside the sample in the pristine state. When a large...
electric field is applied to a cell, the forming process occurs: as shown in Fig. 11(b), a percolating CF is formed and the cell enters the LRS. By applying a suitable external bias, the CF becomes locally ruptured, as indicated in Fig. 11(c). The cell then switches to the HRS (the reset process). The ruptured CFs then reconnect and again form percolating pathways during the set process.

Although Figs. 9 and 10 show that the CFs can have dendritic structures, observing details of these intriguing geometrical structures is difficult with conventional microscopy techniques. The conducting spots in the current maps obtained from the C-AFM experiments therefore correspond to the ends of the percolating CFs. Note that the CF images obtained from the C-AFM and SEM are two-dimensional (2-D) projections of what is actually a three-dimensional (3-D) structure. In addition, TEM images are averaged images of the CF through the sample thickness onto a 2-D plane, so the details of the 3-D structures still cannot be easily investigated systematically. The detailed geometry and the fractal nature of the CFs have been investigated using scaling theory, which is described in Secs. IV C 2 and IV C 3.

b. Temperature, area, and thickness dependence. To better understand the electronic property of a CF, the temperature dependence of the resistance of the LRS and HRS states has been measured experimentally. Figure 12(a) shows the temperature dependence in a pristine state, the HRS, and the LRS of Pt/NiO/Pt cells, which shows unipolar switching. The strong increase of the conductance of the HRS and the pristine state with a temperature increase indicates that both the HRS and the pristine state should be insulating states. In contrast, the conductance of the LRS decreased slightly as the temperature increased, indicating that the LRS is a metallic state, i.e., CFs should exhibit metallic behavior.

Studies have frequently reported that the conductance of a device in the LRS is not dependent on the size of electrodes. Figure 12(b) shows the area–size dependence of Pt/NiO/Pt samples. Note that the LRS of the samples is weakly dependent on the electrode area, which is due to the local nature of the CFs, i.e., the area of the conductive spot is independent of the size of the electrodes. Suppose that a single CF is formed in the LRS; the conductance is not expected to vary with the size of the electrodes. Note that this size-independent behavior of unipolar switching indicates the potential for RRAM to overcome the scaling limits of charge-based memory devices.

However, a few studies have reported that the conductance of the LRS sometimes weakly decreases as the area of the electrodes decreases. This phenomenon has been attributed to the generation of multiple CFs, even though the size of the CF does not depend on the area of the electrodes, multiple CFs can be generated throughout the surface of the oxide. Therefore, if the number of CFs increases with the
the conductivity of the LRS will also weakly depend on the area of the electrodes. In the HRS, the conductivity depends on the electrode area, as shown in Fig. 12(b), which indicates that the local nature of the current pathways is relatively unimportant because the CFs are ruptured. Also, some studies have reported that the conductivity of the HRS is not dependent on the area of the electrodes, which suggests that the current mainly flows in localized regions, even in the HRS. This electrode size-independent behavior might be possible when the HRS is formed with a nearly percolating channel but with a weak CF at the end near an interface, as described in the unified picture of RS in Sec. VB2.

Figure 13 shows the film thickness dependence of the forming, reset, and set processes for the Pt/Fe$_2$O$_3$/Pt cells. The forming voltage is approximately proportional to the thickness of the film. As the forming process has been shown to be triggered by dielectric breakdown with thermal assistance, the applied electric field should be approximately close to the threshold for dielectric breakdown. In the pristine state, the electric field is nearly uniformly distributed throughout the sample to induce dielectric breakdown; therefore, the voltage can be expected to increase in proportion to the film thickness, as shown in Fig. 13(a). However, during the reset and set processes, the rupture and formation of the CFs occur locally, as shown in Fig. 11. Then, the set and reset voltages are required to reconnect and rupture the small local region of CFs, so they would not depend on the film thickness. The lack of dependence of the set and reset voltages on the film thickness is clearly demonstrated in Fig. 13(b).

2. Forming and set processes via soft dielectric breakdown

a. Generation of oxygen vacancies and O$_2$ molecules. In general, as-grown oxide cells are highly insulating and do not exhibit RS. To create RS devices, electroforming is required by applying a large external bias in a controlled manner. This process typically results in the generation of oxygen vacancies, creating CFs inside the oxide cells, as shown schematically in Fig. 11. The electroforming process in unipolar switching is known to be triggered by the soft dielectric breakdown, which is a controlled dielectric breakdown limited by the compliance current. This dielectric breakdown supplies thermal energy to the ions in an oxide cell. Thus, when a large electric field is applied to a pristine oxide cell, oxygen ions move to the anodic interface because they are negatively charged. Oxygen vacancies are widely accepted to be created at the anodic interface via the following electrolytic reaction:

$$O_O \leftrightarrow \frac{1}{2} O_2 + V_O^{-} + 2e^{-}. \quad (3)$$

We follow the notations of Kroger and Vink, and thus $O_O$ denotes an oxygen ion in a regular lattice and $V_O^{-}$ indicates an oxygen vacancy.
Morphological deformations sometimes accompany this process since O₂ gas molecules tend to escape from the surface when negatively charged oxygen ions are attracted toward the top anode. For example, in Pt/TiO₂/Pt bipolar switching cells, the gas eventually erupted from the TiO₂ layer and became trapped as bubbles in the space between the TiO₂ and top electrode layers. Notably, this process can be reversible. A positive bias induces gas bubbles at the top electrode, whereas a negative bias leads to shrinkage of the gas bubbles. When the forming was performed with opposite polarity, the O₂ gas became trapped in the space between the TiO₂ and bottom electrode layers.

The growth behaviors of the CFs during the forming process might differ depending on the dominant force for oxygen vacancy migration. If the electric field effect is the dominant force, positively charged oxygen vacancies are expected to migrate toward the cathode and accumulate there. This process leads to CFs composed of oxygen vacancies accumulating near the cathode and subsequently growing throughout the cell. The CFs generated via this mechanism have been proposed to have a conical structure, which has been observed in TiO₂. This CF growth process is schematically shown in Fig. 14(a). However, if the Soret force (see Sec. III A 3 b) dominates, oxygen vacancies are more likely to move perpendicularly to the electric field. These then form CF pathways that are rich in oxygen vacancies. In this case, the shape of the CFs is expected to be more cylindrical, as shown in Fig. 14(b).

b. The microscopic nature of conducting filaments. The forming process generates localized CFs in an oxide sample. The origin of the local change in conductance is closely related to changes in the chemical stoichiometry of the oxide. The detailed mechanism for how the oxygen stoichiometry will change the local conductance will vary from one material to another: i.e., it should be closely related to the phase diagram of the composing oxide. For example, in the complex phase diagram of titanium oxide, many stoichiometric line phases exist, denoted by TiₙO₂₋ₙ₋₁ (n: integer), close to the TiO₂ phase. These line phases are termed Magnéli phases and are known to have metallic properties. In such a case, the formation energy of the Magnéli phases should be lower than that of the other solid solution phases, so line phases such as these can easily form during the RS. Many other oxides do not have such line phases. In these cases, one may naturally assume that the conducting phases formed during RS should be oxygen-deficient solid solutions.

Formation of the Magnéli phases during RS of TiO₂ cells was directly studied using in situ TEM. In the pristine state, the whole of the oxide was a stoichiometric TiO₂ phase. However, Kwon et al. observed that percolating CFs composed of a Magnéli phase were generated in the TiO₂ layer following the forming process. Therefore, the formation of a CF in the TiO₂ cell appears to be related to formation of the particular phase inside the titanium oxide.

A CF can also be generated by the formation of a metal-rich or metallic region. Park et al. investigated the changes of the microscopic structure of NiO before and after the forming process using a cross-sectional high-angle annular dark-field scanning TEM. This work suggested that the metal-rich region observed in the defective area could result in a local increase in the conductance by creating metallic filaments, which form conducting pathways. Yao et al. also observed similar behavior in an SiO₂ device. Using in situ TEM experiments, they observed an increase of semimetallic silicon nanocrystals during the forming process, which may account for the increased conductance.

The formation of a metal-rich region in NiO is consistent with the results of ab initio calculations of the electronic band structure. Figure 15 shows the calculated partial density of states (PDOS) of Ni atoms using the local density approximation with an on-site Coulomb interaction (LDA+U) simulation. Figure 15(a) shows the crystal structure of NiO; Ni₈₁O₆₄ was used in the simulations. Figure 15(b) shows the PDOS of Ni atoms calculated in the Ni₈₁O₆₄ structure with no vacancies. Figures 15(c), 15(d), 15(e), and 15(f) show the PDOS with 1, 2, 3, and 4 nearest-neighbor oxygen vacancies, respectively. These data show that the presence of oxygen vacancies leads to impurity states near the Fermi level. As the number of oxygen vacancies increases, more impurity states should appear, which makes the system more conductive.

These experimental and theoretical results indicate that main microscopic origin of the local conductivity change in most oxides should be the change in oxygen stoichiometry. When the arrangement of the oxygen ions deviates from that of the ideal stoichiometric sample, or when oxygen vacancies accumulate in a particular region, the conductivity locally increases. During the forming process, such changes can occur non-uniformly and suddenly produce CFs that percolate through the sample. Such a collective behavior is
shown schematically in the bottom panel of Fig. 11. On the other hand, when the atomic structure of a local region returns to the original configuration, the corresponding resistivity returns to the original state in a process that should be closely related to the rupturing CFs.

3. The reset process via rupturing of conducting filaments

When an electric current is localized in a nanometer-scale CF formed during either a forming or set process, the Joule heating produced from the current can raise the local temperature up to several hundred Kelvins. For example, in a Pt/NiO/Pt unipolar switching cell, the maximum temperature in the vicinity of the CF could reach higher than 900 K with a bias of 2 V applied to the electrodes. Note that this temperature is comparable to the processing (such as sintering as annealing) temperature of many oxides, wherein the samples become more homogeneous via solid-state diffusion processes.

When an electric field is applied to a CF, the generated Joule heating can provide the oxygen vacancies with sufficient thermal energy to overcome the energetic barrier for hopping. In this situation, two microscopic forces act perpendicularly to the electric field: the Soret force and the Fick force (see Secs. III A 3 b and III A 3 c). The temperature gradient attracts the oxygen vacancies toward the CF (the Soret force), and the density gradient repels them from the CF (the Fick force). Strukov et al. suggested that the Fick force dominates during the reset process and the CF becomes partially ruptured. Thus, to decrease the free energy by increasing the entropy term, the oxygen vacancies will move from the higher density region of the CFs to the surrounding lower density region, thus rupturing the CFs.

Note that the rupturing process will not make all sections of the CFs disappear at once and enter the pristine state. During the reset process, current will flow into weaker links of CFs, near which local temperatures rise significantly. The rupturing then occurs at the hottest spot, i.e., near the weakest link of the CFs. As shown in Fig. 11, after the rupturing process, the sample does not go back to the pristine state but to the HRS, when a significant portion of the CFs remains intact. The main difference between the HRS and LRS is whether the largest cluster of CFs can form a percolating conduction channel. When a section of the CFs breaks, the current flow across the sample drops suddenly and the sample goes into the HRS.

During the rupturing process, the temperature distribution inside the CFs should play an important role. Russo et al. simulated the CF dissolution mechanism due to Joule heating during the reset process. Figure 16(a) shows a schematic diagram of the simulated geometry. The NiO film (shown by the gray region) was sandwiched by the top and bottom electrodes (shown in the blue regions). The CF (the red cylinder) at the center of the oxide had a radius \( \phi \). Figure 16(b) shows the simulated temperature distributions at applied biases of \( A = 0.53 \) V, \( B = 0.78 \) V, \( C = 0.85 \) V, and \( D = 0.87 \) V during the voltage sweep shown in Fig. 16(c). The black circle symbols denote experimentally extracted temperature data, which are in good agreement with the simulated data shown by the red circles. In Fig. 16(b), the center of the CF has the highest temperature; therefore, it is more like to be ruptured. The green curves correspond to the boundary of the CF. These simulated data clearly show that the rupturing should occur locally due to the nonuniform temperature distribution inside the CF during the reset process.
Note that the localized Joule heating during the reset process also results in nonlinear transport behavior. As described in Sec. III B 1 b, the LRS exhibits metallic behavior. The *I–V* curve for the LRS is therefore ohmic at low voltages, although the *I–V* curve generally becomes nonlinear (i.e., a downward concave) at biases close to the reset voltage. This nonlinear increase in the resistance of the CF is also due to Joule heating. As shown in Fig. 16(b), the local temperature of the CF becomes higher near the reset voltage. Due to the temperature-dependent resistance behavior of a metal, the resistance of the higher temperature region will increase, which explains the nonlinear *I–V* relationships.

C. Bipolar switching

1. What causes polarity-dependent operation?

In comparison to unipolar switching, bipolar switching has polarity-dependence; if the set (reset) process occurs at a positive voltage, the reset (set) process then occurs at a negative voltage, as described in Figs. 4(b) and 4(c). This indicates that some charged species reacting under the electric field exists in an oxide, which induces a resistance change in the material during bipolar switching. Through extensive studies on the basic mechanisms of bipolar switching, two representative charged species are assumed to play the most important roles in inducing bipolar switching: oxygen vacancies and electrons. Secs. III C 1 a and III C 1 b explain how oxygen vacancies and electrons affect the polarity-dependent RS behaviors.

Let us digress a little bit. Note that numerous recent studies have been carried out on the production of two resistance states by changing the orientation direction of dipoles (i.e., polarization) in so-called ferroelectric tunnel junction (FTJ) devices. However, the basic operation of FTJ devices is different from that of RRAM: it does not need a forming process and compliance current. In addition, the resistance changes in FTJ devices are mainly due to the electronic structural changes near the junction and not due to the inhomogeneity, which is different from the main issues of this review. As numerous good review articles have dealt with FTJs, we will not cover FTJs and the associated resistance changes in this review.

**a. Movement of oxygen vacancies.** An oxygen vacancy is a point defect, which is an object in thermodynamic equilibrium. In other words, oxygen vacancies should always exist in any oxide at a finite temperature due to entropic disorder. Oxygen vacancies can be created in numerous ways, but in most cases, they are known to be positively charged defects. For example, many oxygen vacancies can be created via the forming process as explained in Sec. III B 2 a. During the forming process, an electrolytic reaction, Eq. (3), occurs, and positively charged oxygen vacancies are created at the anodic interface. Alternatively, the oxygen vacancies in single-phase oxide materials can be created by partial substitution with dopants, e.g., Y-doped ZrO₂, Gd-doped CeO₂, and Ca-doped BiFeO₃. Due to charge neutrality, the creation of oxygen vacancies always yields electrons to compensate the hole carriers introduced by the doping species, so oxygen vacancies become positively charged.

Since oxygen vacancies have charge, they are forced to move under an external electric field, and depending on the situation, their detailed motions can vary. When the oxygen vacancies are formed in a solid with a periodic potential, their motion under an electric field is described by a simple thermally assisted hopping process, as explained in Fig. 7(a). For SrTiO₃ single crystals, high concentrations of oxygen vacancies were found to exist along a network of extended defects, such as dislocation. In this case, dislocation is known to play the role of a fast transport path for oxygen vacancy migration. Therefore, oxygen vacancies can migrate much more easily through this network under an electric field. Grain boundaries and disordered interfaces can also play very important roles for oxygen vacancy migration.

**b. Trap/detrap of electrons.** Electronic transport in oxides is sometimes dominated by the trapping and detrapping of electrons by defects. These processes are usually described by three conduction models: Poole–Frenkel emission, space–charge limited current, and trap-assisted tunneling. Some review articles and books have described these mechanisms, and the descriptions are summarized as follows. (1) The Poole–Frenkel emission model is basically related to the thermal fluctuations of trapped electrons. The random thermal fluctuations give the trapped electron in localized states enough energy to escape its localized state. Once the electrons transfer to the conduction band, they can move

FIG. 16. Thermal dissolution model. (a) Schematic of the model: φ is the initial diameter of the CF. (b) The calculated temperature maps at different voltages (A, B, C, D) in the voltage sweep (c). In (c), the circular points and the curve denote experimentally extracted and calculated data, respectively. Reproduced with permission from Russo et al., IEEE Trans. Electron Devices 56, 193 (2009). Copyright 2009 IEEE.
through the crystal before becoming trapped by another localized state. The $I–V$ curves behave as $I \sim V^{1/2}$. (2) The space–charge limited current model is considered when the external charge carriers injected from an electrode are more dominant than the free carriers inside the film.273–275 At low voltage, the electrical conduction of the film is dominated by the thermally generated free carriers inside the films over the injected charge carriers. In this case, the $I–V$ curve has a linear relationship as $I \sim V^n$, $n > 1$. At a larger applied voltage, electrons are fully trapped in defects, and finally, excess electrons are able to flow freely again and a linear relationship, $I–V$, is recovered. (3) The trap-assisted-tunneling model is considered when the electrons can be tunneled from trap to trap via hopping.276,277 The electrons are tunneled from cathode to traps (or from traps to anode) in oxides. Like the Poole–Frenkel emission model, the trapped electrons in traps inside the oxides are emitted to the conduction band by random thermal fluctuations. The overall conduction is dominated by tunneling between the cathode (or anode) and the traps in oxides. The $I–V$ curves behave as $\ln(I/V^2) \sim 1/V$. These three conduction mechanisms can be experimentally examined by fitting the $I–V$ curves.

Note that different oxide materials exhibit different trapping and detrapping mechanisms. For ITO/LaAlO$_3$/SrTiO$_3$,269 SrTiO$_3$/YBa$_2$Cu$_3$O$_7$,278 and Cr$_2$O$_3$ films,279 the Poole–Frenkel emission is reported to be the dominant conduction mechanism. For Pr$_{0.7}$Ca$_{0.3}$MnO$_3$,280 La$_{0.7}$Ca$_{0.3}$MnO$_3$,281,282 La$_{0.5}$Sr$_{0.5}$CoO$_{3}$,283 and TiO$_2$ films,100 the observed conduction characteristics are reported to exhibit the trap-controlled space–charge-limited current effects. For TiN/HfO$_2$/Pt films,277 a trap-assisted tunneling model has been reported to be valid.

2. Forming process

a. Formation of conducting filaments. Bipolar switching is known to take place near a metal–oxide interface,36,38,41,44,45 so the forming process does not seem to be required at first sight. Indeed, studies have reported that such a process is not necessary for some oxide thin films.36,155,180,263 However, most other studies have shown that the forming process is needed to obtain bipolar switching, just like unipolar switching. Previously, the necessity for the forming process had presented a great deal of confusion.

To obtain some insight, let us consider some experiments conducted with decreasing film thickness. For a typical film, the forming process is necessary to create the CFs and initiate bipolar switching.33,36,37,284 However, when the film thickness is reduced to a few nanometers, bipolar switching can be initiated without the forming process. Yang et al.36 postulated that forming produces the CFs in the bulk TiO$_2$ film and the actual RS occurs at the metal–oxide interface in bipolar switching. To support their idea, they reduced the thickness of a TiO$_2$ film to 4 nm to eliminate the bulk region and effectively leave only the switching interface. They showed that this device exhibited bipolar switching without the forming process; the $I–V$ curves for the as-grown samples were similar to the $I–V$ curves of the HRS obtained in subsequent bipolar switching operations. Also, other similar experiments have been performed relating to forming free bipolar switching materials, and these have included Al$_2$O$_3$ films,155 and HfO$_2$ films being realized when the film thickness was reduced.

This thickness dependence behavior can be understood by Fig. 2. As discussed in Sec. 1B, the generation of the localized CFs is the most efficient way to induce a very large resistance change, with the resistivity change in a very small volume fraction $p$. Therefore, in most bipolar switching materials, the localized CFs are formed via the forming process as in Fig. 2(d).28,37,285,286 The CFs were observed by TEM experiments in bipolar switching devices such as Pt/TaO$_{x}$/TaO$_{2−x}$/Pt, Pt/TaO$_{x}$/Ta and Pt/TiO$_2$/Pt cells.33,151,159,218 However, in the ultrathin films, with thicknesses as low as 4 nm, the microscopic resistivity change occurs near one of the interface layers, as shown in Fig. 2(c). In this case, the effective $p$ value can be quite large, and the forming process is not necessary for bipolar switching.

For some binary oxides, the chemical composition and crystal structure of the CFs have been characterized by using ex situ microscopic techniques.33,159,218 It happens that the CFs formed in bipolar switching are similar to those from unipolar switching; they have metallic behaviors. In a TaO$_x$ device, the CF was revealed to consist of metallic Ta-rich regions surrounded by a stoichiometric and nanocrystalline Ta$_2$O$_5$ from TEM, EELS, and hard X-ray spectromicroscopy.159 The CFs in TiO$_2$ films were also analyzed by X-ray absorption spectromicroscopy and TEM.33 Within the TiO$_2$ matrix, the formation of a Ti$_2$O$_7$ Magnéli phase was found that exhibited metallic properties and ordered planes of oxygen vacancies. By X-ray spectromicroscopy, ~100 nm conducting channels in TiO$_2$ were also investigated.218 The channels were identified to be oxygen-deficient (vacancy-rich) regions compared with the as-grown material.

b. Bipolar switching operations after the forming process. Since most bipolar switching requires the forming process, we will now focus on bipolar switching with the CFs. The resulting bipolar switching process can be schematized as in Fig. 17. Figure 17(a) shows a pristine sample that does not have many oxygen vacancies. By applying an external bias to the pristine sample, localized metallic CFs, colored brown, are generated in most of the bulk region, as shown in Figs. 17(b) and 17(c). The actual bipolar switching operations take place in the gap region between the electrode and the large cluster of CFs. In the LRS, the gap region, colored yellow, has intermediate resistance: i.e., it becomes more conducting compared with the surrounding oxide matrix, but usually less so than the CFs. In the HRS, the gap region of the intermediate resistance state is recovered to that of the original insulating state. The detailed switching mechanisms that induce the reset and set operations depend on a material, which will be discussed in Sec. III C 2c.
Note that both the LRS and HRS can be reached after the bipolar switching forming process, as depicted in Fig. 17. Some studies have reported that cells (e.g., TiO$_2$ nano-crossbar and Ta$_2$O$_5$/TaO$_2$ nanodevices) entered into the LRS (HRS) when a negative (positive) voltage was applied for the forming process, which differs from the unipolar switching forming process, whereby only the LRS is reached after the forming process.

**c. Role of conducting filaments in bipolar switching.** The high conductivity of the CFs localizes the electronic current in nanoscale conducting spots, as observed on the surface of bipolar switching materials using a C-AFM tip. The electrical conduction of the conducting spots in SrTiO$_3$ thin films can be reversibly modulated over several orders of magnitude by the application of an appropriate electric field. In ZnO films, the bipolar switching was observed in epitaxial nanoslands with a diameter of 30 nm. In WO$_3$ films, the bipolar switching occurred at the grain surface region, not at the grain boundary. Indeed, all of the conductance spots observed in the C-AFM experiments should correspond to the cross section of the localized conducting channels in Fig. 17.

The existence of the CFs can also influence numerous physical properties. When the CFs are formed during bipolar switching, conductivity or resistance will not be dependent on the electrode area. Figure 18 shows the electrode-size dependence on the resistance of bipolar Ta$_2$O$_{5-x}$/TaO$_2$-x cells. As shown in the figure, the resistance of both the LRS and HRS is not sensitive to the electrode area. For comparison, the figure also includes the electrode-size dependence on resistance of the NiO cells exhibiting unipolar switching, which originates from the formation and rupture of localized conducting channels. Similar dependence on the electrode area clearly shows that current is localized in the CFs of Ta$_2$O$_{5-x}$/TaO$_2$-x cells. On the other hand, a few reports have shown that the resistance scales with the electrode size. This type of bipolar switching is sometimes called “homogeneous” interface-type switching, indicating that the CF is absent or not sufficiently strong to localize the current. However, as mentioned in Sec. II B 2, we will call this “laterally uniform switching along the interface.” Figure 18 also shows one example of laterally uniform switching along the interface for Nb:SrTiO$_3$ single crystal cells. However, such examples are quite scarce compared to the cases of CF formation.

Since CFs can be formed in both unipolar and most bipolar switching, many similarities should exist between the two RS phenomena. The schematic diagrams for unipolar and bipolar switching are displayed in Figs. 11 and 17, respectively. The HRS clearly has the same configuration with the largest cluster of non-percolating CFs. An important difference exists only in the LRS. During unipolar switching, the soft dielectric breakdown (Joule heating) will generate (rupture) CFs in the gap region between the largest cluster to the remaining electrode. However, during bipolar switching, the CFs remain intact and the electrical connection in the gap region occurs via polarity-dependent processes, as described in Secs. III C 1 and III C 3. Due to the similarity in Figs. 11 and 17, conversion can occur between the unipolar and bipolar switchings, which will be discussed in detail in Sec. V B.

**3. Models for set and reset processes**

As shown in Figs. 17(b) and 17(c), in most bipolar switching, the actual bipolar switching operations take place
in the gap region between the electrode and the large cluster of CFs. Such polarity-dependent operations have been described by three different models. The first two models are described by the motion of oxygen vacancies, and the last one is associated with the electron trap/detrap processes.

a. Growth/shrinkage of a virtual cathode by oxygen vacancy movements. Since the accumulation (depletion) of oxygen vacancies generally increases (decreases) the conductance, the migration of oxygen vacancies under an electric field can induce a polarity-dependent RS. This mechanism is described in the left-hand panel of Fig. 17(d) and is generally called the “virtual cathode” model.\(^{34,252}\) The virtual cathode denotes a region with a high concentration of oxygen vacancies and thus exhibits a higher conductance compared with the surrounding matrix. Although not clearly stated in the literature, the virtual cathode should be closely related to CFs. In comparison with Figs. 17(b)–17(d), the virtual cathode corresponds to the large cluster of the CFs and the regions of intermediate resistance states, colored yellow.

When a virtual cathode is formed, it will grow from the cathode by oxygen vacancy accumulation because the oxygen vacancies are positively charged, as shown in the figure. After the forming process, the size of the gap between the virtual cathode and the anode becomes only a few nanometers wide. Thus, one can control the percolating path of the virtual cathode by applying an external electric field, which induces bipolar switching. Such an operation should depend on the polarity. At a negative bias, the virtual cathode is attracted to an anode, the oxygen vacancy-rich region becomes percolated, and the cell enters the LRS. At a positive bias, the virtual cathode is repelled from the anode, the percolating channel becomes ruptured, and the cell enters the HRS.

b. Modulation of the Schottky barrier by oxygen vacancy movements. The intermediate resistance state in Fig. 17(b) can originate from a Schottky barrier formed in the narrow gap region between the large cluster of the CFs and the electrode. Note that the Schottky barrier will be formed depending on the choice of electrode. With the \(n\)-type semiconductor, a metal with a small work function, such as Ti, leads to an ohmic interface. On the other hand, a metal with a larger work function, such as Pt or Au, leads to a Schottky barrier at the interface.\(^{137}\) The width of the Schottky barrier is inversely proportional to the square root of the electrode. This results in a narrowing of the Schottky barrier width and changes the HRS into a LRS. Bottom: when a positive voltage is applied at the anode, positively charged oxygen vacancies move towards the electrode. This results in a narrowing of the Schottky barrier width and changes the HRS into a LRS. When a negative voltage is applied at the electrode, negatively charged oxygen vacancies are attracted to the electrode, thereby recovering the original width of the Schottky barrier and inducing HRS-to-LRS switching. When a negative voltage is applied at the anode, oxygen vacancies are repelled from the electrode, thereby recovering the original width of the Schottky barrier and inducing LRS-to-HRS switching.

To provide more quantitative simulations on the Schottky barrier effects, Lee et al.\(^{212}\) developed a semiconductor with a mobile dopant (SMD) model. The SMD model quantitatively considers the effects of a nonuniform distribution of oxygen vacancies on the Schottky barrier width in semiconducting materials. As shown in Fig. 20(a), the SMD model considers a 3-D lattice. Oxygen vacancies (yellow

FIG. 19. Schematics of the set/reset process in bipolar switching. (a) Modulation of the Schottky barrier by oxygen vacancy movements. Top: when a negative voltage is applied to the electrode, positively charged oxygen vacancies move towards the electrode. This results in a narrowing of the Schottky barrier width and changes the HRS into a LRS. Bottom: when a positive voltage is applied at the electrode, oxygen vacancies move away from the electrode. This results in recovery of the original width of the Schottky barrier and induces LRS-to-HRS switching. When a negative voltage is applied at the electrode, electrons are trapped at the defect sites, so many defects remain near the Schottky barrier. This results in narrowing of the Schottky barrier width and induces HRS-to-LRS switching. When a positive voltage is applied to the electrode, electrons escape from the defect sites, so many defects remain near the Schottky barrier. This results in recovery of the original width of the Schottky barrier and induces LRS-to-HRS switching. Reprinted with permission from Seong et al., Adv. Mater. 21, 3754 (2009). Copyright 2009 John Wiley & Sons, Inc. (b) The trap/detrap of electrons. Left: when a positive voltage is applied to the electrode, electrons escape from the defect sites, so many defects remain near the Schottky barrier. This results in recovery of the original width of the Schottky barrier and induces LRS-to-HRS switching. When a negative voltage is applied at the anode, oxygen vacancies are repelled from the electrode, thereby recovering the original width of the Schottky barrier and inducing LRS to HRS switching.
balls) are redistributed due to a hopping motion with thermal assistance in a periodic potential that is modified (i.e., tilted) by an external bias. With a given oxygen-vacancy distribution and an applied voltage, the conduction band energy, or Schottky barrier, can be calculated by self-consistently solving the Poisson equation and the distribution of electrons.

Using the SMD, Lee et al.\textsuperscript{212} found that attracting the oxygen vacancies toward the anode does not always narrow the Schottky barrier width or induce HRS to LRS switching. They found that the opposite can occur when the oxygen vacancy distribution is concentrated near to the anodic interface. They simulated the attraction of oxygen vacancies from the cathode to the anode, as shown in Fig. 20(b). From 0 to 7.5 $\mu$s, the attraction caused the Schottky barrier width to become narrower, as shown in Fig. 20(c). However, further attraction widens the Schottky barrier width, as shown in Fig. 20(d). Using these results, the origin of the coexistence of the cF8 and F8 polarities in one sample was able to be explained (see also Sec. V C 3).

Other effects of oxygen vacancy migration on the modulation of the Schottky barrier have been investigated.\textsuperscript{152,212,250,251} In the reference by Jeon et al.\textsuperscript{250} a change in the Schottky barrier height of an SrRuO$_3$/SrTiO$_3$ (electrode/oxide) junction due to oxygen vacancy migration has been studied using first-principle calculations. They found that the oxygen vacancy inside the SrTiO$_3$ layer lowers the height of the Schottky barrier significantly. However, when the oxygen vacancy is inside the SrRuO$_3$ electrode, the height of the Schottky barrier is comparable to that of an interface with no vacancy. Hur et al.\textsuperscript{152} studied the resistance change of Pt/Ta$_2$O$_5$/TaO$_2$/Pt bipolar switching devices based on the model using a variable Schottky barrier and were able to reproduce the rectifying $I$–$V$ behavior in the HRS and disrupt RS. Jeong et al.\textsuperscript{251} studied the effect of a Helmholtz layer in Pt/TiO$_2$/Pt devices and found that the Schottky barrier height at the interface could be modulated by electrochemical reactions involving oxygen vacancies at an interface between the Pt and TiO$_2$ solid electrolyte.

c. Trap/detrap of electrons. A few models have suggested that the modulation of the Schottky barrier can be obtained by the trapping and detrapping of electrons, instead of oxygen vacancy motions.\textsuperscript{137,154,188} In particular, the claim was made that the trapping and detrapping mechanisms can cause the modulation of the Schottky barrier in defect-rich oxides. Figure 19(b) shows a metal–oxide contact, whereby the Schottky barrier is formed at the interface due to oxygen vacancies.\textsuperscript{188} When a negative voltage is applied to the electrode, the electrons are assumed to be injected into the oxides. They are then trapped at the defect sites and neutralize the oxygen vacancies, making the Schottky barrier wider. As a result, the tunneling current decreases and the device enters the HRS. When a positive voltage is applied to the electrode, electrons are detrapped and the charge of the defects is recovered. The Schottky barrier then narrows again and the device goes into the LRS. Therefore, LRS to HRS and HRS to LRS switching occur when negative and positive voltages are applied to the electrodes, respectively.

Fujii et al.\textsuperscript{172} claimed that trapping and detrapping mechanisms can also induce RS even without affecting the Schottky barrier. In their model, the Schottky barrier is not modulated by a change of the charged defect concentration; instead, the tunneling pathways are controlled by the trapping and detrapping of electrons at trap sites. When a negative voltage is applied at the anode, electrons are injected and captured in the trap sites, which leads to the closure of resonant tunneling pathways. As a result, electrons cannot resonantly tunnel through the Schottky barrier and the cell enters the HRS. When a positive voltage is applied at the anode, trapped electrons are extracted from the trap sites, resulting in the trap sites becoming available as pathways for resonant tunneling through the Schottky barrier, and the cell enters the LRS.

IV. IMPORTANCE OF INHOMOGENEITY AND STATISTICAL PHYSICS APPROACHES TO RESISTIVE SWITCHING

A. Limitations of microscopic approaches: Fluctuations due to inhomogeneity

In Sec. III, the microscopic origins and mechanisms inducing RS phenomena were reviewed. Such microscopic studies enable us to understand why the resistance of a local
region inside a material can change its state during RS operations at the microscopic level. Note that when a sample is inhomogeneous as a metal-insulator composite, as is the case for most RS, its average resistance value may depend on the conductivity value of the metallic region, but it will depend much more on how the highly conducting regions are interconnected. Microscopic studies usually fail to clarify the details of such interconnection, so they do not provide quantitative information on the numerous physical properties of RS.

To obtain further insight, we consider resistor network modeling, which is commonly used in percolation theory in statistical physics, as shown in Fig. 21. Figure 21(a) shows a two-dimensional metal-insulator composite in a real configuration. This can be divided by the coarse-grained lattice denoted by thin lines, as shown in the figure. Cyan color indicates that most of the area in the square is composed of metallic balls, and orange color indicates otherwise, as shown in Fig. 21(b). If two cyan-colored squares are connected, a resistor is placed between them. Then, a resistor network can be obtained, as shown in Fig. 21(b). Figures 21(c) and 21(d) show other configurations of resistor networks with the same number of resistors as that of Fig. 21(b), but with different interconnections. When top and bottom electrodes are attached to the top and bottom edges of the lattices and the total resistances are measured, it is clear that the total resistances of the networks in Figs. 21(b)–21(d) differ significantly from each other due to the different interconnectivity. Therefore, it is important to have theories that can describe how the interconnection between the conducting regions is formed during RS.

One of the most important RS properties, which arises from the form of interconnectivity, is the large fluctuation in transport properties. Nearly, all RS phenomena are accompanied by large fluctuations in transport properties. For example, Lee et al. measured 100 $I$–$V$ curves from seven Pt/NiO/Pt cells, as shown in Fig. 22(a). The red and blue lines correspond to the reset and set processes, respectively. Note that there are wide distributions in $V_{\text{reset}}$ and $V_{\text{set}}$. In some similar studies, the distributions were so large that an overlap...
between $V_{\text{reset}}$ and $V_{\text{set}}$ was observed, so there was no operational voltage window for memory devices.\textsuperscript{75} The resistance values for the LRS and the HRS also vary significantly.\textsuperscript{74,75} Recently, Yoo et al.\textsuperscript{291} systematically investigated the fluctuation between power and resistance for a Pt/NiO$_x$/Pt cell. They applied more than 1900 cycles and collected current-voltage data during both set and reset process. As shown in Fig. 22(b), there were large fluctuations in $V_{\text{reset}}$ and $V_{\text{set}}$, and even larger fluctuations in the corresponding currents. Additionally, they found that their distribution could not be explained in terms of normal or log-normal: wide distributions or long tails were observed for both reset and set switchings. Such large fluctuations in the switching parameters are one of the major problems in RRAM technology.\textsuperscript{89}

These large variations in interconnectivity and the resulting fluctuations are caused partly by the diverse distributions of defects in the pristine state, before the forming process. No material in the real world is free of defects, and numerous kinds of defects are found in nature, including impurities, vacancies, dislocations, and grain boundaries. Even when oxide cells are all fabricated under the same experimental conditions, the detailed distribution of such defects cannot be controlled. So, with an external voltage, the detailed electric field distribution throughout an as-grown oxide will vary from one cell to another. Such an initial variation in electric field distribution will result in different forming, reset, and set processes. Therefore, large fluctuations in the switching parameters become inevitable in all RS phenomena.

The initial variation in the electric-field distribution of an oxide cell leads to fluctuations in the dielectric breakdown voltage. In oxides, oxygen vacancies are the most common defects. When they are located close together, a cluster of oxygen vacancies forms, and this chemically nonstoichiometric object behaves as a conductor. Under an external electric field, a very large local electric field can be generated near the defect clusters, similar to a lightning rod. Such local electric fields can significantly reduce the dielectric breakdown voltage, $V_{\text{forming}}$. For example, the dielectric strength of high purity SiO$_2$ is 470–670 MV/m.\textsuperscript{292} However, most SiO$_2$ cells have a voltage of only $\sim$10 MV/m during the RS forming process.\textsuperscript{116} The magnitude of the reduction in breakdown voltage depends on the location, size, and electrical properties of the oxygen vacancies in the pristine state. Therefore, a random initial defect distribution can give rise to wide fluctuations in $V_{\text{forming}}$.

Additionally, more oxygen vacancies are created during the forming process, and the conducting defect clusters form in an unpredictable way. It has frequently been observed that the electrolysis reaction, Equation (3), occurs at the anodic interface.\textsuperscript{54} Thus, a large amount of oxygen vacancies are created. Under an electric field, these oxygen vacancies migrate to form a cluster, which results in CFs. The location and structure of the newly formed CFs vary significantly, depending on the magnitude of the applied electric field, as well as the initial electric field distribution inside the oxide. It is nearly impossible to predict and control the precise location and structure of a CF in a real sample. This unpredictable nature of structure and location is a general feature of CFs in RS phenomena.\textsuperscript{132} Because the resistance of the LRS depends mainly on the resistance of CFs, the various structures of CFs lead to fluctuations in resistance in the first LRS following the forming process.

After forming, successive RS processes can intensify defect structures and make the interconnections of CFs more inhomogeneous, resulting in fluctuations in switching voltages and resistance values in later processes. As described in Sec. III B 2b, modulation of the local stoichiometry in an oxide induces a localized change in resistance. This modulation also generates defects or inhomogeneous structural configurations during the switching process. Figure 23 shows the creation of such defect structures due to RS.\textsuperscript{234} Figure 23(a) shows cross-sectional TEM images of a pristine NiO cell; there are several grain boundaries, as indicated by the red lines. Following switching, many new grain boundaries are created, as indicated by the yellow lines in Fig. 23(b); furthermore, the form of the grain boundaries is irregular. It follows that the structure becomes more inhomogeneous after RS. This inhomogeneous defect structure changes during successive RS operations, and as a result, the switching parameters also vary persistently.

To understand the physics behind these fluctuations and control them, we need theoretical approaches that can handle the random defect distributions and the structural information of CFs systematically. Note that most microscopic theories do not deal with the detailed electric-field distribution inside the oxide and the structure or the interconnection of CFs. Moreover, defect formation and clustering processes seem to be irregular and random. Due to the randomness of these processes, we can rely on statistical physics...
approaches, especially those based on percolation theories, to understand the large fluctuations and associated phenomena in RS.

### B. Percolation approaches to unipolar switching

Percolation refers to the movement of fluids through porous media, as seen in a coffee percolator. Percolation phenomena are dominated by the properties of connected clusters, specifically passages for fluid movement in the porous media. Therefore, percolation theory deals with the physical and mathematical properties of the connected clusters. One representative percolation model is the bond percolation model suggested by Broadbent and Hammersley in 1957. Following this model, percolation theory has been intensively studied in a variety of fields in mathematics and statistical physics.

The random resistor network model has been widely used to understand the electrical phase transition observed in a metal-insulator mixture. The procedure of modeling from a metal-insulator mixture to a resistor network is depicted in Figs. 21(a) and 21(b). The random resistor network is composed of resistors (or metals) and insulators (unconnected) as shown in Fig. 21(b). The resistance state of each bond is determined randomly: the probability that it is occupied by a resistor is \( p \) and by an insulator is \( 1 - p \). When \( p \) is smaller than a threshold value \( p_c \), there is no percolating path consisting of resistors from one end of the network to the other, so the whole network becomes insulating. In contrast, when \( p \) is larger than \( p_c \), a percolating path is formed and the network becomes conducting. Close to \( p_c \) (\( p > p_c \)), the direct current (DC) conductivity of this network scales as \( \sigma \sim (p - p_c)^\tau \), where \( \tau \) is a conductivity exponent. This scaling exponent exhibits a universal behavior, so it is independent of the details of the material properties. Based on universal scaling behavior, the percolation system can be analyzed using scaling and renormalization theories and the percolating cluster can be described as a fractal structure.

The LRS and HRS of unipolar switching can be viewed conceptually as the appearance and disappearance of a percolating cluster in a network, similar to the random resistor network model, i.e., Fig. 21(b). This simple analogy can enable a deeper understanding of the universal behaviors underlying the observed large fluctuations in unipolar switching. Such a model will also yield a quantitative computer simulation tool for examining the unipolar switching phenomena and methods to control parameters and thereby reduce fluctuations. One important aspect of RS is that the appearance and disappearance of the percolating cluster is considered to be controlled by the external electric field, so it is a time-dependent phenomenon. Additionally, changes between the LRS and the HRS should be repeatable and at least seemingly reversible. Here, “reversible” does not necessarily mean an RS cell returns to the previous state at the microscopic level. In reality, the microscopic state of a CF in one LRS (HRS) could differ from that in another LRS (HRS). Here, “reversible” switching means that the resistance of the cell can be switched back and forth between LRS and HRS. Therefore, a percolation model is needed to deal with the time-dependent and reversible RS processes.

Until 2008, many studies had reported that unipolar switching occurred by generation and rupture of percolating CFs. However, no percolation model could describe the reversible processes in unipolar switching. Note that the random resistor network model by itself is not an appropriate percolation model for RS, because the value of \( p \) is a static parameter and the resistance state of a bond is fixed in time. Therefore, it cannot describe any time-dependent phenomena, including RS. Another percolation model was proposed, called the random fuse network model, where the resistors were replaced by another circuit element, i.e., a fuse. When a current flowing through a fuse is over some threshold value, the fuse blows and its state changes from conducting to insulating. However, this process is irreversible, so it cannot describe the reversibly switchable phenomena. In 2008, to explain time-dependent reversible changes caused by external voltages, Chae et al. used another electrical circuit element instead of resistors or fuses: they used circuit breakers, which can have two bistable states. By making a network of circuit breakers, they developed a new type of percolation model, called the RCB network model.

#### 1. Description of the random circuit breaker network model

Unipolar switching originates from the formation and rupture of CFs, as shown in Fig. 11. Localized CFs are created when the concentration of oxygen vacancies increases. A decrease in the density of oxygen vacancies can induce locally insulating regions. To model these RS phenomena, consider a modeling procedure for oxides similar to that for a resistor network depicted in Figs. 21(a) and 21(b); oxides are divided by a coarse-grained lattice. A square is regarded as a metal if it is sufficiently conductive due to the accumulation of oxygen vacancies, otherwise it is in an insulating state, as shown in the lower panels of Fig. 11. Instead of a normal resistor or a fuse, a circuit breaker, which can be in either an on or off state and permits switching between them, is inserted as a circuit element between the two squares. If the two metal squares are connected, an on-state circuit breaker with resistance \( r_1 \) is inserted, otherwise an off-state circuit breaker with resistance \( r_2 \) is used. Then, the entire oxide cell can be mapped to a network composed of circuit breakers, as shown in Fig. 24(a). This new percolation model is called the RCB network model.

To describe the reversible actions of CFs in the RCB network model, simple switching rules are required to explain how the state of each circuit breaker can be changed between on and off states, as schematically displayed in Fig. 24(b). Switching rules for the changes between on and off states can be imposed by considering the physical mechanisms. First, off/on switching, i.e., off → on, corresponds to the formation of a conducting region due to dielectric breakdown. Thus, off/on switching is assumed to occur when the voltage applied to the off-state circuit breaker, \( \Delta v \), is greater than the threshold value, \( v_{on} \). That is,
In contrast, on/off switching physically corresponds to the recovery of the original insulating state due to Joule heating. Therefore, it is reasonable to assume that the on state changes to the off state by the thermal annealing process. However, the original RCB network model simply assumed that on/off switching occurs when the voltage applied to the on-state circuit breaker $\Delta v$ is greater than another threshold value $v_{\text{off}}$. That is, on-state $\rightarrow$ off-state, when $\Delta v > v_{\text{off}}$.

The switching rule, considering the thermal process, will be reviewed later in Sec. IV B 3 a.

Due to these two switching rules, localized reversible RS can occur. These coarse-grained switching rules are applicable to any RS system regardless of the detailed microscopic mechanisms for RS. Even carbon-based unipolar switching phenomena induced by sp$^3$ (on state)/sp$^3$ (off state) orbital phase transitions, and oxygen-vacancy-mediated RS phenomena, are well described by these switching schemes. Therefore, this model can be material-independent and is focused on the collective behavior of the circuit breakers.

Many natural defects exist inside an as-grown sample. To represent such defects in the pristine state of the simulation, a small fraction of the randomly chosen circuit breakers are assumed to be in the on state as an initial defect. Other kinds of defects can also be used in simulations. For example, a fraction of circuit breakers can be assumed to be randomly eliminated or replaced by conductors, but simulations based on such defects will result in essentially the same RS behaviors. With this initial configuration, an external voltage is applied to the network. The voltage applied to each circuit breaker is calculated by solving Kirchhoff’s equations. Next, the voltage is examined to determine whether it satisfies the switching rules. When off/on or on/off switching occurs in any circuit breaker, the voltage distribution is recalculated and rechecked again using the same switching rules. This calculation should be reiterated until the system reaches a metastable state, where no more switching occurs. Once the iteration process stops at a given value of external bias, another iteration process starts with another bias value. During these simulations, off/on switchings occur in an avalanche-like manner at certain voltages. These correspond to the set and forming processes. During such cases, the iterative process stops when the total current in the circuit-breaker network exceeds the compliance current. Such simple simulations, based on the RCB network model, successfully explain most experimentally observed RS phenomena, as discussed in Sec. IV B 2.

2. Understanding on experimental data

Studying the detailed configuration of RCB networks, obtained by simulations, can help clarify what is happening during the forming, reset, and set processes. For example, Fig. 24(c) shows a sequence of snapshots from a simulation during the forming process. A CF forms via a self-organizing avalanche process. During the simulation, the small seed of the CF (left) grows and finally percolates throughout the network. Note that the simulated shapes of the growing CF are qualitatively similar to the microstructure observed by TEM, shown in the upper panels of Fig. 10.

In particular, a similar dendritic structure of the CF is formed in the simulation. The local nature of the CF generation process explains why unipolar switching has a weak dependence on the size of the electrodes.25,224

---

**FIG. 24.** The random circuit breaker (RCB) network model. (a) Schematic of the RCB network model. The network consists of a circuit breaker, which can be either in $r_1$ (on) or $r_2$ (off) state. (b) Switching rules of a circuit breaker. Off/on (on/off) switchings occur when the applied voltage $\Delta v$ is larger than $v_{\text{on}}$ ($v_{\text{off}}$). (c) Snapshots of the RCB network simulation. First: pristine state of the RCB network. Second: generation of a conducting channel. Third and fourth: percolated state or LRS. Reproduced with permission from Chae et al., Adv. Mater. 20, 1154 (2008). Copyright 2008 John Wiley & Sons, Inc. Comparison between the experiments and simulation results of the RCB network model. Experimental results from TiO$_2$ (d) and NiO (f) thin films. (e)-(g) RCB network simulation results. (d) and (e) Fluctuation in the switching voltages. (f) and (g) Abnormal behavior observed during the reset process. Correspondence between the experiments and the simulations indicates that the various experimental observations in unipolar switching phenomena originate from the collective behavior of the circuit-breaker network. Reprinted with permission from Chae et al., Adv. Mater. 20, 1154 (2008). Copyright 2008 John Wiley & Sons, Inc. and Liu et al., J. Phys. D: Appl. Phys. 42, 015506 (2009). Copyright 2009 IOP Publishing.
Figures 24(d)–24(e) compare the experimental results for polycrystalline TiO$_2$ thin films with the two-dimensional RCB network simulation results. The figures in the left and right panels correspond to experimental and simulation results, respectively. Figure 24(d) shows characteristic I–V curves for unipolar switching measured in numerous TiO$_2$ cells. The pristine state, HRS, and LRS are shown by the green, red, and orange curves, respectively. Although the cells were prepared under the same sample fabrication conditions, significant fluctuations occurred in $V_{\text{forming}}$, $V_{\text{set}}$, and $V_{\text{reset}}$ voltages. Figure 24(e) shows a series of simulated I–V curves for numerous pristine configurations created by choosing the defect circuit breakers randomly, but with the same initial defect fraction. The observed I–V features and fluctuations in switching voltages were reproduced well in the simulations.

Other peculiar RS phenomena, such as irregular reset and multiple switching, have been reported experimentally. For example, in some rare cases of LRS, CFs may be strengthened during the reset process. Such abnormal behavior was reported in NiO thin films by Liu et al., as shown in Fig. 24(f). The current increased abruptly prior to the sudden decrease during the reset process. This strange behavior can be reproduced in RCB simulations, as shown in Fig. 24(g). Although on/off switching is the major transition that occurs during the reset process, off/on switching can also (but rarely) occur for some circuit breakers. This reinforces the CFs and results in a sudden increase in the current, even during the reset process. Another unusual but frequently observed behavior is multiple switching: this phenomenon can also be reproduced and understood using an RCB network model, including multiple reset process (see also Sec. VI D).

Besides explaining the fluctuating feature of the I–V curves, the RCB network model can also explain the distribution of resistance. Kim et al. performed an interesting experiment using two stainless steel balls with diameters of 0.95 cm. The surface of the stainless steel balls was covered in a 5 nm-thick oxide layer. When the two balls were in contact with each other, as shown in the left figure of Fig. 25(a), the contact point (red boxed area) could be approximated to a capacitor-like structure, i.e., electrode-oxide-electrode. However, when sufficient current was applied, an abrupt decrease in resistance occurred, possibly due to the generation of percolating CFs through the oxide layer. These measurements were repeated easily after vibrating the balls and obtaining a new pristine oxide layer. Figure 25(b) shows the experimentally obtained distribution of stainless steel ball contact resistance after applying 1 μA. Two peaks occurred: a high-resistance peak at ~1.2 MΩ (without percolating CF) and a low-resistance peak at ~20 Ω (with percolating CF). The low-resistance peak was well fitted to the log-normal distribution, as shown in the figure. This distribution was confirmed by a three-dimensional RCB network model. Figure 25(c) shows the simulated resistance distribution of the oxide layer with percolating CFs in the LRS. As the figure shows, the RCB network model can be used to explain the electrical contact resistance between the stainless steel balls. All of the results shown in Figs. 24(d)–24(g) and 4–5 demonstrate that the RCB network model is quite successful in describing unipolar switching, especially fluctuation phenomena.
3. Variations in the switching rules of the random circuit breaker network model

a. Temperature-dependent switching rule. As described in Sec. IV B 1, the original RCB network model assumed for simplicity that on/off switching occurs when the voltage applied to the on-state circuit breaker is greater than a threshold value. Later, Chang et al. used a more realistic model, called the “thermal RCB network model,” where the on/off switching rule is governed by temperature instead of voltage. In this model, the electric current and temperature increase due to Joule heating were calculated for each circuit breaker by solving the Laplace and heat transport equations simultaneously. Here, the heat dissipation to the thermal bath from the circuit breaker was also taken into account, as described in Fig. 26(a). If \( T \) is the local temperature of a circuit breaker at a certain moment, on/off switching is dependent on the following switching rule, as shown in Fig. 26(a):

\[
\text{on-state} \rightarrow \text{off-state}, \quad \text{when} \quad T > T_c.
\]

Note that the off/on switching rule is the same as the original one, i.e., based on the applied voltage to the circuit breaker.

This temperature-dependent switching rule gives a clear explanation of the nonlinear behavior of the \( I-V \) curve near \( V_{\text{reset}} \). The inset of Fig. 22(a) shows a nonlinear \( I-V \) curve in a \( \text{TiO}_2 \) cell; the curve deviates from the ohmic line near \( V_{\text{reset}} \). Similar nonlinear \( I-V \) curves have been observed in most unipolar switching materials. This nonlinear behavior can be reproduced by the thermal RCB network model. Figure 26(b) shows a simulated \( I-V \) curve for a reset process using the thermal RCB network model. Figures 26(c)–26(e) correspond to the circuit-breaker configurations obtained from the simulations. The configuration at each numbered point in the \( I-V \) plot is denoted by the same label, and the red shading indicates elevated temperatures. Note that the temperature rise is greatest in the local region near the weakest link in CFs. The Joule heating is insufficient at a low bias, so the \( I-V \) curve remains in a linear regime. Close to the reset, the temperature increases significantly. This higher temperature increases the resistance, which results in a deviation from the ohmic response.

Note that the Joule heating is concentrated at the narrowest region of the CF, so that temperature-dependent on/off switching occurs near the hottest point of the CF network. Only a few circuit breakers are switched off and the percolating conducting cluster becomes ruptured during the reset process. During the set process, the CF becomes reconnected in the ruptured region. In contrast to the forming process, only a few circuit breakers are switched on during the set process and the required voltage \( V_{\text{set}} \) can be lower than \( V_{\text{forming}} \). This local rupturing and reconnecting of CFs explains the observed lack of dependence of \( V_{\text{reset}} \) and \( V_{\text{set}} \) on film thickness, as shown in Fig. 13(b). The temperature-dependent switching rule also explains other RS behaviors quite well, including conversion between unipolar memory and threshold switchings due to variations in temperature and electrode thickness (see Sec. V A 1).

b. Stochastic switching rules. In the original and thermal RCB network models, switching rules are deterministic: if a proper switching condition is satisfied, the state of the circuit breaker will always be changed. It is also worth mentioning that the switching rules can be given in a probabilistic way, i.e., off/on or on/off switching occurs with a given probability. For example, the fluctuation behavior of NiO RRAM cells has been investigated using a percolation model with stochastic switching rules. The NiO layer is approximated by a three-dimensional lattice of small identical cubes. When an external bias is applied to the top and bottom electrodes, the electric field and temperature for the whole lattice can be calculated. According to the calculated electric field and temperature, the state of each cube is updated. First, off/on switching of a cube occurs with probability \( p_1 \) as:

\[
\text{off-state} \rightarrow \text{on-state with probability } p_1 \sim p_{10} \exp \left( -\frac{GT}{\Delta V} \right),
\]

where \( \Delta V \) is voltage applied to the cube, \( G \) is an experimentally determined constant, and \( p_{10} \) is a normalization constant. This probability was introduced to describe the dielectric breakdown. Second, on/off switching of a cube occurs with probability \( p_2 \) as:

\[
\text{on-state} \rightarrow \text{off-state with probability } p_2 \sim p_{20} \exp \left( -\frac{E_a}{k_B T} \right),
\]

where \( E_a \) is the activation energy, \( k_B \) is the Boltzmann constant, and \( p_{20} \) is a normalization constant. This probability...
originates from the Arrhenius equation because thermal dissolution is induced by oxygen vacancy migration. Using these two probabilistic switching rules, the authors reported that they could obtain more realistic and accurate calculation of switching parameters for NiO RRAM cells.

C. Scaling behaviors

Variations in cell resistance and switching voltages are commonly observed in unipolar switchings, as shown in Fig. 22. At first, these variations appear to be random and unpredictable, but careful analysis reveals that they are governed by some simple scaling relations. Moreover, these scaling relations could be material-independent and provide valuable information about the structure of the CFs. In this section, we present some experimental results for such scaling behaviors, and then describe how these scaling relations are related to the fractal structure of the CFs.

1. Scaling behaviors of physical properties

a. Scaling behavior of 1/f noise. Noise is an important feature of all electrical circuits, and the characteristics of noise are of scientific and technological significance. From an applications perspective, a smaller (or minimal) noise for RRAM is highly desirable. Additionally, because noise can have several different origins, analysis of noise can provide valuable information about the electronic transport and the microscopic switching mechanisms. Analyses of the noise in many electronic devices and electrical circuits typically reveal that the power spectral density is inversely proportional to the frequency $f$ in the low-frequency regime, which is termed Flicker noise, or 1/f noise. This 1/f scaling behavior has also been found in cells exhibiting both unipolar and bipolar switchings. Figure 27(a) shows the 1/f noise measured in a Pt/NiO/Pt cell exhibiting unipolar switching. Note that the noise intensity, $S_V$, in the HRS is several orders of magnitude larger than that in the LRS.

There are a number of possible origins of 1/f noise. The observed 1/f noise in unipolar switching should be closely related to the resistance noise in a random resistor network, which consists of a fraction of resistors $p$ and a fraction of insulators $1-p$. When $p$ is larger than a percolation threshold, $p_c$, the resistors form a percolating cluster through the whole network, and an electric current, $I$, can flow. The noise power, $S_R$, is defined as

$$S_R = \frac{S_V}{f^\delta}.$$  

When $p$ is just slightly larger than $p_c$, random resistor network theory expects that

$$R \propto (p - p_c)^{-1},$$

and

$$S_R \propto (p - p_c)^{-\kappa},$$

where $R$ is the total resistance of the random resistor network, and $\beta$ and $\kappa$ are positive exponents. Thus

$$S_R \propto R^{2+\beta},$$

where $\beta + 2 = \kappa/\delta$. It follows that the noise power and resistance will have a scaling relation close to the percolation threshold, with an exponent, $\delta$. Since unipolar switching can be mapped to a network of circuit breakers, it is expected that a similar scaling behavior may be observed.

This idea has been experimentally confirmed, as shown in Fig. 27(b). As expected from random resistor network theory, for the LRS, $S_R \propto R^{2+\beta}$ with the exponent $\delta = 1.8 \pm 0.3$. For the HRS, $S_R/R^2$ becomes very large and is scattered around $10^{-3}$Hz$^{-1}$. The exponent $\delta$ for the HRS decreases with the ambient temperature and appears to saturate at $\delta \sim 1.6$. The RCB network simulation model also

![Figure 27](image-url)
Fig. 28. Scaling behaviors during the reset process. (a) Scaling behavior between \( I_{\text{reset}} \) and \( R_0 \) measured from NiO thin films with a 30 \( \times \) 30 \( \mu \)m\(^2\) electrode area. Two scaling regimes are separated at the cross-over resistance \( R_{co} \). Reproduced with permission from Lee et al., Appl. Phys. Lett. 93, 212105 (2008). Copyright 2008 AIP Publishing. (b) Scaling relation between \( I_c \) and the resistance of semi-continuous metal percolating films. Reprinted with permission from Yagil et al., Phys. Rev. Lett. 69, 1423 (1992). Copyright 1992 American Physical Society.

gives an exponent of \( \alpha \sim 1.5 \). This scaling behavior in \( 1/\tau \) noise indicates that the CFs in unipolar switching can be explained in terms of percolation theory.

Similar \( 1/\tau \) scaling behaviors were also observed in HfO\(_x\) cells exhibiting bipolar switching, as shown in Fig. 27(c). For the LRS, the \( S_f \) spectra always follow \( 1/\tau \) scaling. However, for the HRS, they have a \( 1/\tau \) dependence at low frequency, but there is a crossover from \( 1/\tau \) to \( 1/\tau^2 \) with the increase in \( \tau \). The authors attributed the observed \( 1/\tau^2 \) behavior in the HRS to electron tunneling between the electrode and the nearest trap sites.

b. Scaling behaviors between reset current, third harmonic signal, and resistance in the low-resistance state. A wide variation in \( I-V \) curves is very common in unipolar switching. Figure 22(a) shows 100 \( I-V \) curves from 7 NiO thin films with a 30 \( \times \) 30 \( \mu \)m\(^2\) electrode area. The variation in the reset voltage \( V_{\text{reset}} \) appears random and unpredictable. The inset shows the LRS resistance at low voltage \( R_0 \) and the reset current \( I_{\text{reset}} \). \( R_0 \) and \( I_{\text{reset}} \) also exhibit wide variations, which are seemingly random. However, simple scaling relations were found to explain these fluctuations.

Figure 28(a) shows \( I_{\text{reset}} \) as a function of \( R_0 \) from the \( I-V \) curves that were plotted in Fig. 22(a). As the figure shows, \( I_{\text{reset}} \) follows a simple power law, \( I_{\text{reset}} \propto R_0^{\gamma} \), with a crossover at \( R_0 = R_{co} \): for \( R_{co} > R_0 \), \( \gamma = 1.8 \pm 0.2 \), and for \( R_{co} < R_0 \), \( \gamma = 0.7 \pm 0.1 \). The two different scaling regimes also exhibited different features in the \( I-V \) characteristic curve. When \( R_{co} > R_0 \), both \( V_{\text{reset}} \) and \( I_{\text{reset}} \) decrease as \( R_0 \) increases. However, when \( R_{co} < R_0 \), \( V_{\text{reset}} \) increases but \( I_{\text{reset}} \) decreases as \( R_0 \) increases. Similar power-law behaviors with somewhat different exponents were also reported for CoO thin films with 0.5-mm electrode diameter films \(^{314} \) for \( R_{co} > R_0 \), \( \gamma = 1.2 \pm 0.1 \), and for \( R_{co} < R_0 \), \( \gamma = 0.6 \pm 0.1 \).

Interestingly, similar scaling behavior was observed much earlier in classical percolation systems. Yagil et al. prepared semi-continuous metal percolating films via the evaporation of metals.\(^{99} \) They applied a current to both sides of these films and measured the breakdown current \( I_c \) and the resistance before breakdown. Here, \( I_c \) is defined as the current at which the film underwent the first irreversible resistance change, where hot spots were formed and the melting point of the metallic grain was reached. This process is analogous to the reset phenomena, and hence \( I_c \) is similar to \( I_{\text{reset}} \). They found that \( R_0 \) and \( I_c \) exhibited scaling behavior with \( I_c \propto R_0^{\gamma} \), and that there was a crossover resistance \( R_{co} \sim 2\ \Omega \) as shown in Fig. 28(b). For \( R_0 < 2\ \Omega \), \( \gamma = 1.75 \pm 0.4 \), and for \( R_0 > 2\ \Omega \), \( \gamma = 0.85 \pm 0.2 \). This scaling behavior was attributed to the detailed structure of the narrow bottlenecks, which carried a large current density in the percolating network.\(^{99} \) Therefore, the existence of similar scaling behavior observed in unipolar switching suggests that such power laws originate from the peculiar geometry of CFs, particularly a bottleneck structure.

Along with the relation between \( I_{\text{reset}} \) and \( R_0 \), a scaling behavior was also found between \( R_0 \) and the third harmonic signal. Here, the third harmonic signal \( B_{3f} \) is an electrical response with a tripled frequency \( 3f \), where the input frequency is \( f \). Figure 29 shows the plot for \( B_{3f} \) and \( R_0 \) normalized by crossover values measured in NiO\(_x\), SrTiO\(_3\), FeO\(_y\), and TiO\(_z\) cells.\(^{335} \) They follow a power law, \( B_{3f} \propto R_0^{\lambda} \) with a crossover: for \( R_{co} > R_0 \), \( \lambda = 3.1 \), and for \( R_{co} < R_0 \), \( \lambda = 6.7 \).

The power law relationship between \( B_{3f} \) and \( R_0 \) indicates there should be a simple scaling principle for the structure or interconnectivity of CFs. This is because \( B_{3f} \) and resistance are strongly dependent on the microgeometry of CFs. First, from the random resistor network theory, when an alternating current \( I = I_0 \cos(2\pi ft) \) is applied to both side bars of the network, \( B_{3f} \) is proportional to the fourth moment of

Fig. 29. Scaling plot of \( B_{3f} - R_0 \) for NiO\(_x\), SrTiO\(_3\), FeO\(_y\), and TiO\(_z\). Here, \( B_{3f} \) and \( R_0 \) are normalized by their crossover values. Reprinted with permission from Lee et al., Phys. Rev. Lett. 105, 205701 (2010). Copyright 2010 American Physical Society.
current flowing through all resistors in the resistor network, as follows:

\[ B_{3f} \propto \frac{1}{I_0^2} \sum_n r_n^{1.4} \]  

(4)

where \( r_n \) is the resistance of a resistor labeled \( n \) in the resistor network, and \( i_n \) is the local current flowing through the resistor \( r_n \). Note that because higher moments of the current distribution are much more sensitive to microgeometry details, \( B_{3f} \) can be used as a nondestructive tool for investigating changes in interconnectivity within the resistor network. Second, the total resistance of a resistor network, \( R \), is proportional to the second moment of the current distribution. From the conservation of energy, i.e., \( I^2R = \sum_n i_n^2r_n \), the resistance, \( R \), is given by

\[ R = \frac{1}{I^2} \sum_n r_n i_n^2. \]  

(5)

From Equations (4) and (5), it is clear that both \( B_{3f} \) and total resistance depend on how the conducting resistors are interconnected within the resistor network. Therefore, the scaling behavior shown in Fig. 29 suggests that there should be a certain scaling property in the interconnectivity of CFs. This scaling property was found to be the fractal geometry of the CF, which will be discussed in Secs. IV C 2 and IV C 3.

Note that the scaling exponents for the \( B_{3f} - R_0 \) plot of NiO, SrTiO, FeO, and TiO cells are the same, as shown in Fig. 29. The exponents for the \( I_{\text{reset}} - R_0 \) plot of the four oxides were also found to be the same, indicating that the observed scaling exponents are universal and material-independent. This universality, which is one of the most important concepts in statistical mechanics and solid state physics, denotes a property in which some physical quantities are independent of microscopic details in the system. Therefore, if one physical, observable behavior belongs to a universal class, it is not changed by varying the material.

However, material-independent observable behavior can usually be changed by varying a dimension or the size of a system. This has also been observed in NiO nanowire systems with unipolar switching. Kim et al. fabricated NiO nanowires with diameters of 200, 80, and 30 nm, as shown in Fig. 30(a), and measured the scaling exponents from the \( I_{\text{reset}} - R_0 \) and \( B_{3f} - R_0 \) plots. Note that as the diameter of the nanowire decreases, the oxide approaches a one-dimensional system. Figure 30(b) shows the measured \( I_{\text{reset}} - R_0 \) plot. As the diameter of the nanowire decreases from 200, to 80, to 30 nm, the exponent increases as \( 0.86 \pm 0.04, 1.25 \pm 0.02 \), and \( 1.28 \pm 0.06 \), respectively. This clearly shows that the exponents depend on system dimension or size. Figure 30(c) shows the measured \( B_{3f} - R_0 \) plot. In this case, unlike in the \( I_{\text{reset}} - R_0 \) plot, the scaling exponent is not dependent on the system size but is a constant value of \( 3.25 \pm 0.06 \). Another interesting finding from this nanowire experiment is that there is only one scaling regime, while there are two scaling regimes in thin film experiments.

The existence of the scaling behaviors described above also reveals a very important limitation in microscopic theories. The behaviors do not originate from individual resistance changes in the microscopic parts of an oxide, but from how they cooperatively interact with each other under an electric field, and construct a unique interconnectivity of CFs. Therefore, the origin of these scaling behaviors cannot be understood using microscopic approaches. This limitation is very critical for RRAM applications, because RS phenomena are changes in transport properties and associated fluctuations will limit the performance of RRAM. In contrast, an appropriate scaling theory based on statistical physics can reveal the physics behind these scaling behaviors.
2. Experimental observations of fractal structures in percolating conducting filaments

As mentioned earlier, it is important to investigate the structural properties of CFs to understand and control unipolar switching phenomena. Some experimental works have addressed the shape of CFs. It was previously assumed that CFs were cylindrical, but further studies involving high resolution TEM revealed conical CF structures. The overall shape of CFs could be varied depending on the microscopic forces during CF formation, as shown in Figs. 30(a) and 30(b). However, as in classical percolation systems, the interconnection inside the 3-dimensional CF is a more important geometrical factor affecting transport properties.

It is quite difficult to investigate the exact 3-dimensional structure of CFs because most experimental imaging techniques (such as TEM and SEM) use 2-dimensional projection. However, some reports have revealed the complex nature of CF geometry. Figure 23(b) shows that a complex network of grain boundaries was formed in a NiO cell following the switching process, which is believed to form conducting pathways. As shown in Fig. 10, a dendritic structure of the CF reportedly appeared during the forming process in a ZnO cell. As shown in Fig. 31(a), a similar dendritic structure was also reported in a Pt/CuO/Pt cell based on SEM studies with in-plane geometry. Note that multiple CFs were formed following the forming process. These results indicate that CFs might not be formed in simple cylinders or cones, but may instead have complex dendritic-like structures.

Growth in dendritic structures has been reported in numerous systems in nature. The results of an electro-deposition experiment, which involves the same mechanism as electro-chemical metallization. When an external bias is applied to an (initially clean) sample, Ag+ ions migrate from the anode toward the cathode. The light-colored areas at either side of the figure are metal electrodes (Ag and Ni). Note that dendritic clusters of Ag particles started to form near an electrode but grew into a metal electrodes (Ag and Ni). When an external bias is applied to the initial clean sample (upper), the dendritic Ag-structure is formed (lower). This is the basic mechanism for an electro-chemical metallization cell. Reprinted with permission from Fujiiwa et al., Jpn. J. Appl. Phys., Part 1 47, 6266 (2008). Copyright 2008 IOP Publishing. (b) Electrodeposition experiment. Reprinted with permission from Kozicki et al., Phys. Rev. Lett. 52, 1033 (1984). Copyright 1984 American Physical Society. (c) A Lichtenberg figure induced by electric breakdown. Reprinted with permission from Niemeyer et al., J. Non-Cryst. Solids 352, 567 (2006). Copyright 2006 Elsevier B.V. (d) A Zn-dendritic structure formed during a zinc electrodeposition experiment. Reprinted with permission from L. M. Sander, Nature 322, 789 (1986). Copyright 1986 Macmillan Publishers.


The mathematical properties of these dendritic patterns were extensively investigated in the 1980s and 1990s. Metal ion aggregation during electro-deposition has been studied using the diffusion limited aggregation model. Earlier studies demonstrated that the dendrite patterns exhibit a fractal dimensionality.

A fractal is a mathematical object exhibiting self-similar or scale-invariant patterns. Fractals are distinguished from ordinary integer-dimensional objects in that they have a fractional dimension. For example, doubling the diameter of a circle increases the area by a factor of 2², and doubling the diameter of a sphere increases the volume by a factor of 2³. However, if the length of one side of a fractal is doubled, the volume of the fractal is increased by 2²D, where D is termed a fractal dimension and has a non-integer value. For example, the fractal dimension of the radially grown Zn-dendritic structure on the 2-dimensional plate shown in Fig. 31(d) is ~1.7. In this case, D can be determined from the following scaling relation:

\[ L(t) \propto t^D, \]  

where \( t \) is the radius of an imaginary circle, the center of which coincides with the center of the dendrite, and \( L(t) \) is the total length of the dendritic structure within the imaginary circle.

Interestingly, although these patterns are formed by different mechanisms, the dielectric breakdown pattern in two dimensions, shown in Fig. 31(c), has the same fractal dimension (of ~1.7) and electro-deposition pattern. This is not a coincidence, because the pattern-formation processes induced by electro-deposition and by 2-dimensional dielectric breakdown belong to the same universal growth process, i.e., Laplacian growth. Note that CF formation in unipolar switching occurs via aggregation of oxygen vacancies, the migration of which is triggered by dielectric breakdown. Therefore, it is reasonable to expect that the dendritic
structure of the CF formed after the forming process may be inherited from the Laplacian growth.

3. Scaling theory based on fractal structure

a. Derivation of scaling exponents. A scaling theory, of how the CFs are interconnected inside the sample, has been developed to understand the scaling properties in unipolar switching systems. In this section, we review this scaling theory for CF. As in other dendrite systems, it can be assumed that the total length, $L$, of all branches inside a circle of radius, $t$, is proportional to $t^{D_f}$, as follows from Equation (6). Thus, $n_t \sim dL(t)/dt \sim t^{D_f-1}$ and $\rho_t \sim n_t/\pi t^2 \sim t^{D_f-3}$, where $n_t$ is the number of branches and $\rho_t$ is the density of the branches at a given distance, $t$, from the center. For CFs in unipolar switching, $n_t$ can be defined as the number of branches passing through a cross-section, where the distance from the top surface is $t$ and $\rho_t$ is the density of the branches, as shown in Fig. 32(a). The diffusion limited aggregation pattern may be inherent to the geometry of the current paths in the CF. Then, $n_t$ and $\rho_t$ can be expressed as

$$n_t \sim (c_0 + t)^{D_f-1},$$

and

$$\rho_t \sim (c_0 + t)^{D_f-3},$$

where $c_0$ is a positive constant. The constant, $c_0$, is inserted because $n_0$ should be non-zero at $t = 0$ because there will be a finite bottleneck area. Here, $c_0 \propto s_0$, where $s_0$ is the lateral size of the bottleneck, because $s_0 \sim \sqrt{A_0} = \sqrt{n_0/\rho_0} \sim c_0$, where $A_0$ is the spot area at $t = 0$. Therefore, $s_0$ can be substituted for $c_0$ in the derivation of scaling relations.

Using the scaling relations given in Equations (7) and (8), Lee et al. calculated the resistance, $R_0$, and the third harmonic signal, $B_{3f}$. From Equations (4) and (5), $R_0$ and $B_{3f}$ are proportional to the second and fourth powers of the current distribution, respectively,

$$R_0 \propto \sum_m \left( \frac{i_m}{I_{tot}} \right)^2,$$  \hspace{1cm} (9)

and

$$B_{3f} \propto \sum_m \left( \frac{i_m}{I_{tot}} \right)^4,$$  \hspace{1cm} (10)

where the index $m$ denotes each conducting bond, and $I_{tot}$ is the total current passing through the network. If the electric current is assumed to be uniformly divided among $n_t$ branches at the cross-section $t$, then, from Equation (9) the resistance is given by

$$R_0 \propto \sum_{i=0}^{s_0} \left( \frac{1}{n_i} \right)^2 \cdot n_t = \sum_{i=0}^{s_0} \frac{1}{(s_0 + t)^{D_f-1}} \sim \begin{cases} s_0^{-D_f+1} & \text{for small } s_0, \\ s_0^{-D_f+2} & \text{for large } s_0. \end{cases}$$

In the same way, from Equation (10), $B_{3f}$ becomes

$$B_{3f} \propto \sum_{i=0}^{s_0} \left( \frac{1}{n_i} \right)^4 \cdot n_t \sim \begin{cases} s_0^{-3D_f+3} & \text{for small } s_0, \\ s_0^{-3D_f+4} & \text{for large } s_0. \end{cases}$$

Note that the above scaling relations, i.e., Equations (11) and (12), are divided into two regimes depending on the size of $s_0$, as shown in Fig. 32(b). This can explain the observed two scaling regimes, as shown in Figs. 28 and 29.

The reset current $I_{reset}$ can also be evaluated. The rupturing of CFs occurs at a bottleneck ($t = 0$), because the bottleneck will be the hottest point in the CF network due to the largest current density. If $i_{reset}$ is defined as the threshold current to induce transition to an insulating state at a given branch of the network, then the reset current, $I_{reset}$, is given by

$$I_{reset} \approx i_{reset} n_0 \sim i_{reset} s_0^{D_f-1}.$$  \hspace{1cm} (13)

Using Equations (11)–(13), with the assumption that $i_{reset}$ does not depend on $s_0$, we may calculate the scaling exponents between the quantities $R_0$, $B_{3f}$, and $I_{reset}$. In fact, however, $i_{reset}$ depends on $s_0$ because of the Joule heating dissipation process. Taking into account Joule heating, the reset current is given by
TABLE III. Comparison between measured and calculated exponents $\lambda$, $\gamma$, and $\eta$ for low (subscript l) and high (subscript h) resistance regimes for various experiments. Data for (a), (b), and (c) were extracted with permission from Lee et al., Phys. Rev. Lett. 105, 205701 (2010); Wang et al., J. Phys. D: Appl. Phys. 43, 385105 (2010); and Kim et al., Appl. Phys. Lett. 104, 023513 (2014), respectively.

<table>
<thead>
<tr>
<th>Theoretical formula</th>
<th>$\lambda_l$</th>
<th>$\lambda_h$</th>
<th>$\gamma_l$</th>
<th>$\gamma_h$</th>
<th>$\eta_l$</th>
<th>$\eta_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{2l} \sim R_{c0}^{-3}$</td>
<td>3</td>
<td>6.7 ± 0.3 (6.7)</td>
<td>3.1 ± 0.2 (3)</td>
<td>1.8 ± 0.2 (1.8)</td>
<td>0.7 ± 0.1 (0.61)</td>
<td>6.7 ± 0.3 (6.7)</td>
</tr>
<tr>
<td>$B_{2h} \sim R_{c0}^{-3}$</td>
<td>3</td>
<td>3.25 ± 0.06 (3)</td>
<td>3.25 ± 0.06 (3)</td>
<td>1.2 ± 0.1 (1.2)</td>
<td>0.6 ± 0.1 (0.59)</td>
<td>3.25 ± 0.06 (3)</td>
</tr>
<tr>
<td>$I_{Rl} \sim R_{c0}^{-\gamma}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_{Rh} \sim R_{c0}^{-\gamma}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_{Rl} \sim B_{2l}^{\lambda}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_{Rh} \sim B_{2h}^{\lambda}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where $\phi = 1.650.01$.\textsuperscript{235} The scaling exponents can be calculated by using Equations (11)–(14) and are listed in the first row of Table III. In the table, subscript $l$ and $h$ indicate the cases for $R_0 < R_{c0}$ and $R_0 > R_{c0}$, respectively.

b. Comparison with experimental results. Table III also contains the exponent data from the experiments described in Sec. IV C 1 b. The exponents in the second row were measured from experiments involving NiO thin films.\textsuperscript{125} These experimental values were also confirmed in studies involving NiO$\textsubscript{x}$, SrTiO$\textsubscript{y}$, FeO$\textsubscript{z}$, and TiO$\textsubscript{w}$ cells.\textsuperscript{235} With a fractal dimension of 2.54, which is the fractal dimension of usual 3-dimensional diffusion-limited aggregation models, the experimentally measured exponents agree well with the theoretical values, as presented in the table. The data in the third row were measured from CoO thin films with a 0.5 mm electrode diameter.\textsuperscript{114} In this case, when the fractal dimension is 2.96, two measured exponents correspond to the theoretical values quite well. It is not clear why the fractal dimension in this case is close to 3 but the difference may originate from the system size difference: the size of the NiO layer is $30 \mu m \times 30 \mu m \times 80 nm$ and that of CoO is $\sim 0.5 mm$-diameter $\times 150 nm$.

For a nanowire, the effective dimension can be varied by the ratio between diameter and length. If the ratio approaches zero, the nanowire should behave as a one-dimensional object. In contrast, the effective dimension should increase with an increase in diameter. As mentioned earlier, Kim et al.\textsuperscript{316} performed scaling experiments using nanowires with various diameters. In the fourth, fifth, and sixth rows of Table III come from nanowires with 200, 80, and 30 nm diameters, respectively. For all these cases, the exponent for the $B_{3l}R_{0}$ plot is 3.25, which is in reasonable agreement with the theoretical value of 3. The fractal dimension estimated from the $I_{\text{reset}}R_{0}$ plot decreases as the diameter of the nanowire decreases. This should originate from the confinement of CFs: for a very narrow nanowire, the CFs should be confined to narrow cylindrical space, which forces them to have a one-dimension-like shape. This may be the reason for the fractal dimension coming close to 1 when the diameter is reduced. This experiment suggests that the scaling theory could be applicable to a device with width down at least to 30 nm. However in this size limit, the CF in a device should behave as an almost one-dimensional object.

V. CONVERSION BETWEEN RESISTIVE SWITCHING TYPES AND THE IMPORTANCE OF STATISTICAL PHYSICS

A. Conversion between memory and threshold switchings

As introduced in Sec. II B, there are two types of unipolar switching whose $I$–$V$ curves are symmetrical about the applied bias, i.e., memory and threshold switching. Memory switching has two meta-stable resistance states, as shown in Fig. 4(a), and its mechanism is explained in Sec. III B. In contrast to memory switching, threshold switching has only one stable resistance state without an external bias, as shown in Fig. 4(d).\textsuperscript{1,2,6,181} Memory and threshold switching are seemingly quite different and have usually been considered independently; the connection between them has rarely been discussed.

However, several recent experiments have reported the co-existence of both memory and threshold switching in the same NiO cells\textsuperscript{108–112} and HfO$\textsubscript{2}$ cells.\textsuperscript{327} The conversion between memory and threshold switching has also been observed using several methods: namely, by modulating the film stoichiometry,\textsuperscript{26,112} by varying the thickness of the electrode,\textsuperscript{109} by changing the ambient temperature,\textsuperscript{110} and by controlling the compliance current.\textsuperscript{111,112,327} Such coexistence and conversion of memory and threshold switching indicate that the basic mechanisms of these two switching types should be related. Therefore, a deeper understanding of the conversion mechanism for these switching types will provide a better understanding of both phenomena. In Secs. V A 1 and V A 2, we deal mainly with the mechanism of the coexistence of memory and threshold switching, based on the stability of CFs and the associated percolation model, known as the thermal RCB network model.

1. Experimental observation of the conversion between memory and threshold switchings

Since the early 2000s, interesting observations of the coexistence of memory and threshold switching in the same material have been reported. Seo et al. observed that memory...
swapping in NiO films could be turned into threshold switching when the oxygen content was increased during growth. By using Rutherford back-scattering and X-ray diffraction methods, a higher concentration of nickel vacancies was found in the NiO samples exhibiting memory switching, compared with the samples showing threshold switching. The authors insisted that a lack of nickel vacancies resulted in the absence of charge trapping sites near the Fermi level, making the LRS unstable, leading to threshold switching.

Similar observations of both memory and threshold switching were also made by changing the thickness of the bottom electrode. Chang et al. fabricated Pt/NiO/Pt cells with various thicknesses of bottom electrode. For all cells, the fabrication conditions and the thickness of the NiO layers were kept constant. These cells, however, exhibited different switching types depending on the thickness of the bottom electrode, $t_{BE}$: devices with a smaller $t_{BE}$ exhibited threshold switching, whereas devices with a larger $t_{BE}$ exhibited memory switching. More experiments on this conversion were carried out on the same sample using thermal cycling. Figures 33(a)–33(d) show the switching type change when the $I–V$ curves of the NiO devices were measured at low (118 K) → high (300 K) → low (80 K) → high (300 K) temperatures. The $I–V$ curves at both low and high temperatures consistently represented memory and threshold switching, respectively. This result indicates the importance of thermal effects in conversion between memory and threshold switching.

In addition, the compliance current was found to be an important factor in controlling the switching type between memory and threshold switching. With Ni/NiO core-shell nanowires, He et al. found that threshold switching occurred when the compliance current was set to 10 µA, while memory switching was observed when the compliance current was set to 100 µA. With HfO$_2$ cells, Saura et al. also showed that both threshold and memory switching could be obtained by using different values of the compliance current during the forming process. They found that a cell exhibits threshold switching when the compliance current is lower than 100 µA, but that memory switching occurs with a higher compliance current value. With NiO thin films, however, Peng et al. observed the opposite behavior; they reported that memory switching was observed at a compliance current of 1 mA, but that it turned into threshold switching at 10 mA.

2. Mechanisms for the conversion between memory and threshold switchings: The thermal random circuit breaker network model

Two groups have attempted to attribute the conversion of switching type to film stoichiometry. Peng et al. performed first-principles calculations on NiO to calculate the electronic band structure depending on the chemical stoichiometry. Figure 34(a) shows the electronic band structure of stoichiometric Ni$_{32}$O$_{32}$, which involves the upper/lower Hubbard band and the oxygen 2p band. They argued that a high electric field could facilitate transfer of electrons from the oxygen 2p valence band into the upper Hubbard conduction band. The resulting metallic Ni is responsible for the formation of CFs in memory switching, as shown in Fig. 34(b). Since the energy gap between the upper Hubbard band and the oxygen 2p band is sufficiently large, electrons cannot return with the small electric field, resulting in stable CFs. Figure 34(c) shows the electronic band structure of non-stoichiometric Ni$_{33}$O$_{32}$ with nickel vacancies. The oxygen 2p band is split and an additional band is formed just above the Fermi energy. The charge transfer is not stable since the Fermi level is between the lower and higher oxygen 2p bands. They argued that a low electric field mainly excites the charge fluctuations between the two oxygen 2p bands, and a stable metallic filament cannot be maintained, giving rise to threshold switching, as shown in Fig. 34(d). From this scenario, they claimed that the modification of band structure holds significance in the volatility of RS.

A more convincing scenario was formulated to explain the conversion between memory and threshold switching in terms of the thermal stability of CFs due to Joule heating. As described in Sec. III B 3, Joule heating provides energy for oxygen vacancies to be released from the oxygen-vacancy
chance, a circuit breaker in the on state could be switched back to the off state due to a higher temperature. This indicates that the stability of CFs is affected by their local temperature: at low temperature the CF is stable, whereas the CF becomes unstable at high temperature. Therefore, a stable CF can be formed during the set process at low ambient temperature, which results in memory switching. However, when the surrounding temperature is sufficiently high, a stable CF can rarely be formed during the set process, which leads to threshold switching.

By using the thermal RCB network model, it was shown that thermal stability of CFs is the basis of the temperature cycling experiment displayed in Figs. 33(a)–33(d). In contrast to the original RCB network model, Chang et al. assumed that a circuit breaker in the on state could be switched back to the off state when the temperature exceeded a threshold value, \( T_c \) (see Sec. IV B 3 a for a detailed explanation of the thermal RCB model and its switching rules). Figure 33(e) shows the simulated \( I-V \) curves from the thermal RCB network model. When the thermal bath temperature, \( T_b = 0.3 \ T_c \), the obtained \( I-V \) curves exhibited memory switching, as shown in the left panel of Fig. 33(e). In this case, circuit breakers in the on state formed a percolating cluster inside the network, following the set process. When \( T_b = 0.75 \ T_c \), threshold switching was predicted, as shown in the right panel of Fig. 33(e). In this case, a percolating cluster of on-state circuit breakers could not be formed because the on-state circuit breakers generated during the set process were immediately switched to the off state due to a higher temperature. These results demonstrate that the conversion between threshold and memory switching can result from thermal instabilities in CFs in the LRS.

The conversion experiment investigating the impact of changing the thickness of the bottom electrode, introduced in Sec. V A 1, can also be explained by the same origin as the temperature cycling experiment. Kim et al. investigated the origin of the conversion by using the three-dimensional thermal RCB network model. They performed voltage-sweeping simulations using the circuit-breaker networks depicted in Figs. 35(a) and 35(c) with different bottom electrode thicknesses \( t_{BE} = 0.2 \ \mu m \) and \( t_{BE} = 0.01 \ \mu m \), respectively. They found that immediately after the forming process, the oxide temperature increased to \( \sim 500 \ \text{K} \) for the device with \( t_{BE} = 0.2 \ \mu m \), as shown in Fig. 35(a). In contrast, with the same conditions, the oxide temperature for the device with \( t_{BE} = 0.01 \ \mu m \) increased to \( \sim 600 \ \text{K} \), as shown in Fig. 35(c). This temperature difference originates from the fact that heat sinks more effectively through the thicker bottom electrode. Therefore, a percolating cluster of on-state circuit breakers can be formed in the sample with the thicker bottom electrode, which results in memory switching, as shown in the \( I-V \) curve in Fig. 35(b). However, for the sample with the thinner bottom electrode, threshold switching appeared after the forming process because the percolating cluster cannot be formed due to high temperature.

Finally, further studies are needed to clarify the reason that threshold switching appears at both low and high values of compliance current. Some researchers have claimed that during the set process with a low compliance current value, the produced Joule heating is insufficient to move the oxygen vacancy to form nonvolatile CFs. Thus, only a temporary electrical pathway is formed (as also described in Sec. III A 3 b), and threshold switching occurs. On the other hand, Peng et al. claimed that a large amount of Joule heat generated during the set process increase the concentration of oxygen ions in the cell. So, with a higher compliance, the Joule heat disturbs the formation of an oxygen-vacancy cluster and threshold switching occurs.

B. Conversion between unipolar and bipolar switchings

In Secs. III B and III C, the mechanisms of unipolar and bipolar switching are discussed, respectively. In unipolar switching, percolating channels of CFs are generated by aggregation of the oxygen vacancies via soft dielectric breakdown during the forming process. On the other hand, in the bipolar switching forming process, it is widely accepted for various oxides that CFs are generated but fail to form percolating channels between the electrodes. The gap between the CF and the electrodes remains insulating or semiconducting, but its resistance can be altered by the motion of oxygen vacancies or electrons. Note that some reports, such as of TaO cells, claim that a metallic CF percolates through the system in the LRS of bipolar switching. However, such exceptional cases have been very limited. Therefore, in this section, we will deal with the case of both the LRS and HRS of bipolar switching showing...
nonmetallic behavior, which is typical for cells exhibiting a coexistence of unipolar and bipolar switching.

If the mechanisms of unipolar and bipolar switching are independent of each other, it should not be possible to initiate conversion between the two RS modes in one sample. However, there have been numerous experimental reports on mode conversion between unipolar and bipolar switching.100–107,238,328 These experiments suggest that both switching phenomena share some common mechanisms. Specifically, if unipolar and bipolar switching have a common percolation nature due to the existence of CFs, it would be possible to develop a statistical model for bipolar switching by extending the RCB network model for unipolar switching. Therefore, careful investigation of conversion between switching modes could provide further insight into the underlying physics of both unipolar and bipolar switching systems and provide greater understanding of the fluctuations in bipolar switching. In Secs. V B 1–V B 4, we review studies on the coexistence of bipolar and unipolar switching, as well as on conversion between the two modes, and discuss the mechanisms of conversion. Then, we describe how to develop a percolation model for bipolar switching devices involving CFs.

Note that unipolar and bipolar switching have an LRS and HRS. It is difficult to distinguish between the two resistance states, i.e., the LRS and HRS, in each of the unipolar and bipolar switching modes. Therefore, there is a need to introduce new notations that will enable distinction. In Secs. V B 1–V B 4, we use the following new abbreviations: uLRS and uHRS to denote the LRS and HRS in unipolar switching, respectively, and bLRS and bHRS to denote the LRS and HRS in bipolar switching, respectively.

1. Experimental observations of the conversion between unipolar and bipolar switchings

It has been reported that many oxide materials, including NiO,103,329,330 TiO2,101,102,105,107 SrTiO3,100,238,331,332 ZnO,328,333–335 Mn:Ba0.7Sr0.3TiO3,104 HfOx336,337 TaOx338 and BiFeO3,339 exhibit the coexistence of bipolar and unipolar switching in the same device. Figures 36(a) and 36(b) show an example of the coexistence of unipolar and bipolar switching observed in a single Pt/SrTiOx/Pt cell.100 As denoted by the blue dashed line in Fig. 36(a), a pristine SrTiOx cell exhibited bipolar switching following the initial forming process; set and reset processes occurred at different polarities. After the application of a large positive voltage (red solid curve), the bipolar switching changed to unipolar switching. Then, the set and reset processes occurred with the same voltage polarity. Figure 36(b) shows that, during repetitive switching operations, a sudden mode change in the SrTiOx sample can occur from unipolar to bipolar switching. Although this seemingly reversible switching-mode conversion has been observed in a few oxide materials,100,104 these experiments indicate that the observed unipolar and bipolar switching modes are not independent phenomena, but are in fact closely related.

From the literature, there are a couple of important control parameters that can induce the conversion between unipolar and bipolar switching. The first is the magnitude of the compliance current. During the forming process, a large compliance current usually leads to unipolar switching, whereas a smaller compliance current leads to bipolar switching.102,104 In such cases, conversion from bipolar to unipolar switching can also occur by increasing the compliance current during the set process. The second important control parameter is the voltage-sweep range. Shen et al.104 observed switching-mode conversion when controlling the voltage-sweep range using Pt/Mn:Ba0.7Sr0.3TiO3/Pt cells. When they initially swept the voltage from 0 to 10 V, the cell exhibited unipolar switching, with a reset voltage of 0.7 V and a set voltage of 4 V. Specifically, the uHRS changed to the uLRS above 4 V in unipolar switching. On the other hand, when the cell was in the uHRS, if the
voltage-sweep range was reduced to 2 V, less than the unipolar switching set voltage, the mode changed to bipolar switching. The application of a (negative) voltage resulted in a cell entering a more (less) resistive state. This bipolar switching mode can revert to unipolar switching when the voltage-sweep range is returned to the original range. In this way, unipolar and bipolar switching can be obtained repeatedly by controlling the voltage-sweep range.

The aforementioned experimental reports provide us with some important insight into the relationship between unipolar and bipolar switching. First, more electrical power, i.e., a larger compliance current or switching bias range, is required for unipolar compared with bipolar switching.100,104,105,107,238,328–332,338,339 This implies that Joule heating plays a more important role in unipolar switching. Second, the uHRS could be closely related to the bipolar switching operation for the following reasons. (i) Existence of non-percolating CF clusters is common in the uHRS and both bipolar switching resistance states. As described in Sec. III B 1 a, locally ruptured CFs are formed in the uHRS. And, as explained in Sec. III C 2, the forming process is required for bipolar switching to generate similar non-percolating CFs.44 (ii) The temperature dependence of samples in the bHRS and bLRS is typically that of a nonconductor, as in the uHRS. Figure 37 shows the temperature dependence of all resistance states during mode conversions of Ba0.7Sr0.3TiO3 thin film.104 Only the uLRS resistance increases with temperature, showing a metallic behavior. For the resistances of all other states, resistance decreases with temperature, i.e., it behaves as an insulator or a semiconductor. Similar temperature dependence was also observed in SrTiO3 cells100 and ZnO cells.333 These temperature-dependence experiments support the hypothesis that CFs percolate through the sample only in the uLRS. The nonmetallic behaviors of uHRS suggest that this state might be related to bipolar switching where both resistance states have nonmetallic temperature dependence.

2. Unified description of unipolar and bipolar switchings

The experimental observations discussed in Sec. V B 1 can be explained qualitatively by using the following unified picture of bipolar and unipolar switching, illustrated in Fig. 38. By applying a proper external bias to the pristine sample, shown in Fig. 38(a), either bipolar or unipolar switching can be generated. The resulting switching modes depend on the electrical power dissipated in the oxide cell during the forming process. When a relatively small electrical power is applied during the forming process, bipolar switching can be obtained. In this switching mode, extended CFs are formed but fail to generate a percolating channel, as shown in Figs. 38(b) and 38(c). On the other hand, when a larger electric power is applied, unipolar switching can be obtained, as shown in Fig. 38(d). The resulting CFs form a percolating channel between the top and bottom electrodes, and the sample goes into the uLRS. All of these forming processes are irreversible: after forming, the original pristine state cannot be recovered by applying an external voltage.

Figures 38(b) and 38(c), and the corresponding red lines, represent the bLRS, bHRS, and the operations of bipolar switching, respectively, which are consistent with Fig. 17 in

![Fig. 36. Conversion between unipolar and bipolar switchings in a Pt/SrTiO3/Pt cell. (a) I–V curves showing mode conversion from bipolar to unipolar switching. (b) I–V curves showing mode conversion from unipolar to bipolar switching. Simulated I–V curves of reversible RS-type changes between bipolar and unipolar switchings using the interface-modified RCB network model. (c) RS-type change from bipolar to unipolar switching. (d) RS-type change from unipolar to bipolar switching. Reproduced with permission from Lee et al., Appl. Phys. Lett. 98, 033502 (2011). Copyright 2011 AIP Publishing.](image)

![Fig. 37. Temperature dependence of the HRS and LRS for both bipolar and unipolar switching. Reproduced with permission from Shen et al., J. Appl. Phys. 107, 094506 (2010). Copyright 2010 AIP Publishing.](image)
This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 147.46.19.34 On: Tue, 13 Oct 2015 06:35:15

Fig. 38. Unified picture of the unipolar and bipolar switchings. (a) Pristine state. Green and blue areas represent an oxide and the electrodes, respectively. If the electric field is not sufficiently high during the forming process, an incomplete percolating path, denoted by a brown path in the oxide, is formed as (b) and (c). Then, the bipolar switching mode can be initiated. bLRS (b) and bHRS (c) of bipolar switching can be reversibly obtained by the application of the opposite external bias. The yellow dashed path shown in (b) denotes the region where the local resistance is reduced compared with the surrounding oxide. The resistance of the yellow region is larger than the brown path. If sufficiently high electrical power is applied during the forming process, a percolating path is formed as (d), which results in the uHRS. When a large amount of Joule heating are applied to the percolating path, it ruptures as (b) and (c), which are also the uHRS. The uLRS and uHRS can also be reversibly switched to each other by the application of a high electric field with the same polarity.

Sec. III C 2. Note that the bLRS (Fig. 38(b)) and bHRS (Fig. 38(c)) have different properties in the region between the extended CF and the electrode. The two states can be reversibly obtained by applying an external bias of the opposite polarity. The channel near one of the interfaces in the bLRS, marked by a yellow dashed line, has a lower resistance than the surrounding area. To distinguish this channel from the ordinary CF, it is termed a weakly conducting filament (wCF), since it typically shows markedly lower conductance than the metallic CF, and usually, with some exceptions, has semiconducting temperature dependence. 159 On the other hand, such a wCF disappears in the bHRS. A wCF can be formed under the electric field as a result of various microscopic origins, explained in Sec. III C 3. For example, the wCF can originate from a reduction in the Schottky barrier width due to the migration of oxygen vacancies, 44 or from the opening of a tunneling path due to the extraction of trapped electrons. 172 It should be noted that the other parts of CFs, connected to the wCF in the bulk, are not removed during bipolar switching, so these invariant channels can be interpreted as a virtual cathode. 44

Figures 38(c) and 38(d), and the corresponding blue lines, represent the uHRS, uLRS, and the operations of unipolar switching, respectively, that are consistent with Fig. 11 in Sec. III B 1 a. The main difference between Figs. 38(c) and 38(d) is whether the CF channel becomes percolated or not. When the CF percolates through the entire oxide cell, corresponding to the uLRS, its resistance is significantly smaller than that of the uHRS. If a large current flows through this percolating CF, significant Joule heating is produced and the filament is ruptured, leading to the uHRS as shown in Fig. 38(c). These percolating and ruptured states can be reversibly switched via set and reset processes with the same polarity, which corresponds to unipolar switching. Note that the configuration of Fig. 38(b) can also be reached during unipolar switching. Since the generation of the wCF depends on the polarity, it is logical that the state in Fig. 38(b) could be reached, depending on the polarity when the percolating CF is ruptured. However, since the resistance difference between Figs. 38(b) and 38(d) is typically very large, the configuration of Figs. 38(b) and 38(c) can be recognized as the uHRS.

Within this unified picture, the mode conversion experiments presented in Sec. V B 1 can be easily understood. These are possible because bipolar and unipolar switching share very similar CF configurations. When the thermal annealing process due to Joule heating locally ruptures the percolating CF in the uLRS, the resulting state should belong to one of the bipolar switching configurations. In this situation, if a relatively small electrical power (by setting a small compliance current or sweeping over a smaller voltage range) is applied, a fully percolating CF is not formed, but a polarity-dependent physical process, such as oxygen vacancy movement or an electron trap, can occur. So, bipolar switching can be obtained. On the other hand, if greater electrical power is applied (by setting a large compliance current or sweeping over a larger voltage range) to the material, a percolating CF is re-generated, and the operating mode changes to unipolar switching. Therefore, studies on the conversion between bipolar and unipolar switching can provide us with an important testing ground for a unified picture of switching.

3. Percolation model for bipolar switching: The interface-modified random circuit breaker network model

A unified picture, Fig. 38, was quantitatively described by a statistical percolation model, called the interface-modified RCB network model. 100 In the original RCB network model, locally conducting and insulating regions of the oxide are modeled as the on and off states of a circuit breaker, respectively, as shown in Fig. 24(a). In the interface-modified RCB network model, there should also be on (metallic) and off (non-metallic) states. The difference is that the off state is divided into two sub-states: off1 and off2, where the resistance of off1 is lower than that of off2. The reason that off1 and off2 states are introduced is to describe the forming and disapparing states of the wCF, respectively, in the bLRS (Fig. 38(b)) and bHRS (Fig. 38(c)).

In the interface-modified RCB network model, off1 and off2 states are assumed to be located only near the interface, as shown in Fig. 39(a). 100 In principle, two types of off state can exist in any location, either inside the bulk or near the interface. However, numerous experimental studies have reported that the weakest part of the CF, i.e., the most
probable location for rupture, might be near the interface and that the physical phenomena that underpin bipolar switching occur mainly near the interface. Based on these observations, for simplicity in the simulation, the off1 and off2 states are assumed to occur only in the interface region. Therefore, in the interface region, three resistance states in the circuit breaker should be considered: on, off1, and off2, whose resistances are denoted by $r_{l_{\text{bulk}}}$, $r_{l_{\text{int}}}$, and $r_{l_{\text{int}}}$ ($r_{l_{\text{bulk}}} < r_{l_{\text{int}}} < r_{l_{\text{int}}}$), respectively. For the bulk region, there are two resistance states, as in the original RCB network model: on and off states, whose resistances are denoted by $r_{l_{\text{bulk}}}$ and $r_{l_{\text{bulk}}}$, respectively.

The switching rules for the circuit breakers at the interface and in the bulk are presented schematically in Figs. 39(b) and 39(c), respectively. For the bulk circuit-breakers, the switching rules are the same as those of the original RCB network model, as described in Sec. IV B 1. For circuit breakers near the interface, two types of switching rule are used depending on the polarity: (i) polarity-dependent switching rules between $r_{l_{\text{int}}}$ (off2) and $r_{l_{\text{int}}}$ (off1) causing bipolar switching; and (ii) polarity-independent switching rules between $r_{l_{\text{int}}}$ or $r_{l_{\text{int}}}$ and $r_{l_{\text{bulk}}}$ (on) causing unipolar switching. Detailed switching rules can be found in Ref. 100.

Because these polarity-dependent switching rules are given without specifying detailed microscopic mechanisms, the interface-modified RCB network model was applied to investigate the statistical features of any bipolar switching phenomena with extended CFs. Except for the switching rules for interfacial circuit breakers, the simulation methods for the interface-modified RCB network model are essentially identical to those for the original RCB network model, as described in Sec. IV B 1.

One of the advantages of simulations based on the RCB network model is that they can provide us with detailed configurations of what occurs during RS processes. Figure 39(d) shows how the forming process occurs in bipolar switching. Initially, wCFs develop in the interfacial region. Since they can provide a non-uniform voltage distribution inside the bulk, like a lightning rod, a soft dielectric breakdown is more likely to occur from the wCFs, and the CFs will be generated from that point. The resulting CF does not percolate through the sample, so bipolar switching occurs. As shown in Fig. 39(e), when the compliance current is small, the sample can be switched between the bHRS and bLRS. In the bipolar switching mode, the extended CFs in the bulk change little and most of the change occurs in the gap between the CF and the electrode. Note that the
configurations in the left and right sides of Fig. 39(e) correspond to Figs. 38(b) and 38(c), respectively. When the compliance current becomes large, the circuit breakers in the gap can adopt the on state (i.e., with $I^{\text{bulk}}$) and then the CF will percolate through the sample, as shown in the left picture of Fig. 39(f). Then, unipolar switching can occur. Note that the configurations in the left and right sides of Fig. 39(f) correspond to Figs. 38(d) and 38(c), respectively.

Experimentally observed mode conversions between bipolar and unipolar switching can be reproduced in interface-modified RCB network model simulations. Figures 36(c) and 36(d) show simulated $I-V$ curves from the interface-modified RCB network model. After a forming process with a low compliance current value, bipolar switching occurred, as denoted by the blue dashed curve in Fig. 36(c). When a higher compliance current was set, the switching changed to unipolar switching, as denoted by the purple curve in Fig. 36(c). Sometimes, this unipolar switching mode could be converted back to a bipolar switching mode, as shown in Fig. 36(d). When a lower compliance current was set after this transition, bipolar switching operated reliably again. These simulation results explain well the experimental results in Figs. 36(a) and 36(b).

Several scanning probe microscopy studies have reported local conducting spots appearing on oxide surfaces in the bLRS. When a conductive area on a thermally reduced SrTiO$_3$ bipolar switching device was scanned by C-AFM, it was observed that nanoscale conductive spots were distributed unevenly on the surface. The simulations also explain why local conducting spots can be observed in the bLRS. They should originate from the CFs in the bulk region, generated in the forming process. The region where the CF is nearly touching the oxide surface should have a higher local conductance than its surrounding region.

The simulations also explain why the forming process was not required in some very thin films during bipolar switching. Most pristine cells have a very large resistance due to their thick bulk region, which is not suitable for switching operations. This is the reason that the forming process is needed. With the extended CFs generated inside the bulk during the forming process, the effective thickness for RS is reduced significantly, as can be seen in Figs. 38(b) and 38(c). Then, bipolar switching can take place at the interface region. Therefore, when the film thickness is reduced to less than several nanometers, bipolar switching can occur without the extended CF. Such forming-free bipolar switching devices were realized in 4-nm-thick TiO$_2$ cells.

4. Artificial control of unipolar to bipolar switching by inserting an interfacial layer

Secs. V B 1–V B 3 discussed mode conversion between unipolar and bipolar switching in the same sample, induced by controlling compliance current value or voltage range. The mechanism of the mode conversions can be explained by the precise control of CFs near the interface layer. However, the boundary between the two layers is not apparent for a real capacitor-type cell since it is usually prepared by depositing a single oxide layer. Note that the local microstructure of the extended CF inside the bulk layer can be varied by adjusting several factors: the defect configuration of the pristine state; the applied voltage; the compliance current; and so on. Therefore, it is sometimes difficult to choose the correct experimental parameters that make it possible to observe mode conversion between unipolar and bipolar switching in a single-layer cell.

Systematic studies using two-layer cells, namely, by artificially inserting an insulating layer between the electrode and the original single layer, were performed to provide new insights into mode conversion. Yoo et al. fabricated Pt/TaO$_x$/Pt cells and artificially inserted an insulating Ta$_2$O$_5$ layer between the top electrode and the TaO$_x$. Figures 40(a) and 40(b) show the cross-sectional TEM image of Pt/TaO$_x$/Pt and Pt/Ta$_2$O$_5$/TaO$_x$/Pt cells, respectively. As shown in Fig. 40(c), the Pt/TaO$_x$/Pt cell demonstrated unipolar switching after the forming process. However, when the Ta$_2$O$_5$ was inserted, the resulting cell underwent bipolar switching, as shown in Fig. 40(d). In addition, by increasing the Ta$_2$O$_5$ layer thickness, the forming voltage and the current value of the bLRS could be controlled systematically.

These interesting results can also be explained by the interface-modified RCB network model, whose results are shown in Figs. 40(e)–40(h). When there is no interface layer, as shown in Fig. 40(e), unipolar switching occurs, as shown in Fig. 40(g). The cell enters the uLRS when the CFs percolate through the sample, while it enters the uHRS when the percolating CF becomes disconnected. In contrast, when the Ta$_2$O$_5$ interface layer is inserted as shown in Fig. 40(f), CFs consisting of on-state circuit breakers do not percolate the interfacial Ta$_2$O$_5$ layer, but exist only inside the bulk region. Then, bipolar switching occurs, as shown in Fig. 40(h). This work demonstrates that RS characteristics can be controlled by inserting an artificial layer.

Similar research using double-layer cells has been performed by other researchers, since these cells have many advantages for use in practical RRAM applications, namely, high scalability, fast switching, high endurance, good retention, and low power consumption. Lee et al. at Samsung Electronics attributed such high performance to a double-layer Pt/Ta$_2$O$_5$/TaO$_x$/Pt structure, where the TaO$_x$ layer acts as a uniform conductor and the CF develops in the Ta$_2$O$_5$ layer, as shown in Fig. 41(a). Note that this structure differs from that in Fig. 40(f), i.e., the interpretation by Yoo et al. of the RS mode in the Pt/Ta$_2$O$_5$/TaO$_x$/Pt cell changes.

Recently, the Samsung research team revisited the Ta$_2$O$_5$/TaO$_x$ bilayer structure by using in-situ scanning TEM. From cross-sectional high-angle annular dark-field scanning TEM images, they produced pseudo-color maps of the structural evolution in a TaO$_x$ layer, as displayed in Fig. 41(b). Yellow and blue colors represent the pseudo-color maps of the material's electron density in the TaO$_x$ layer. After the set process (left panel), the bright and yellow regions appeared, indicating the formation of conductive percolation paths in both the Ta$_2$O$_5$ and TaO$_x$ layer. After the reset process (right panel), the bright and yellow regions partially disappeared near the top (i.e., in the Ta$_2$O$_5$ layer).
but many CFs remained in the TaOx layer. The TEM result indicated that the CFs were formed inside both the TaOx and the Ta2O5 layer, as schematically shown in Fig. 41(c). A major change, inducing bipolar switching, occurs in the insulating Ta2O5 layer. The CFs inside the Ta2O5 layer are connected (top panel) and ruptured (bottom panel) for the set and reset processes, respectively, while the CFs inside the TaOx layer do not undergo many changes. These recent TEM findings support the explanation of the interface-modified RCB network model in Fig. 40.

C. Conversion between figure-of-eight and counter-figure-of-eight polarities in bipolar switching

As introduced in Sec. II B 2, one puzzling phenomenon in bipolar switching is that there are two types of hysteretic $I–V$ curve: F8 type and cF8 type. With F8, as shown in Fig. 4(b), the $I–V$ curves exhibit HRS-to-LRS and LRS-to-HRS changes following the application of positive and negative voltages, respectively. In contrast, as shown in Fig. 4(c), the cF8 $I–V$ curve shows LRS-to-HRS and HRS-to-LRS changes at positive and negative voltages, respectively.

Numerous mechanisms have been proposed to explain the different switching polarities.

For F8-type bipolar switching, an electronic model based on trapping and detrapping of electrons at fixed defect sites, as described in Sec. III C 3 c, has been commonly used. In this model, a Schottky barrier is also formed at the anodic interface. Application of a negative voltage to the anode injects electrons into the oxides. Then, trapped electrons at the defect sites neutralize the oxygen vacancies, which widens the Schottky barrier and the device enters the HRS. On the other hand, a positive voltage makes electrons detrapped, and then, the width of the Schottky barrier becomes narrower and the device goes into the LRS. Thus, the corresponding $I–V$ curve should be the F8-type.

The coexistence of both types of bipolar switching in the same sample has also been reported experimentally in numerous materials, such as Fe-doped SrTiO3, Sr2TiO4, SrTiO3, and Nb0.007O3,348 and TiO2.37,290,349 Here, the switching-type coexistence is not attributed to the simple change of the grounded reference electrode. For example, Fig. 42(a) shows $I–V$ curves obtained using a symmetric Pt/TiOx/Pt device structure. A Schottky barrier was formed at both top and bottom electrode/oxide interfaces due to the large work function of the Pt. The authors observed that both polarity modes could be switched by applying an external bias. Figure 42(b) shows another $I–V$ curve of Au/Sr2TiO4/Nb:SnTiO3,346 where both cF8 and F8 bipolar switching can exist in the same cell. A Schottky barrier was formed only at the Au/Sr2TiO4 interface, and the other electrode/oxide interface exhibited ohmic behavior. This cell also exhibited coexistence of the two polarities that could be reversibly switched by an external bias.
between F8/cF8 switching types. To date, three possible mechanisms have been proposed, based on the role of oxygen vacancy migration in bipolar switching. The first model was suggested to explain the conversion observed in symmetric devices, e.g., Pt/TiO\textsubscript{x}/Pt cells, where a Schottky barrier is formed at both top and bottom interfaces (see Sec. VC1).\textsuperscript{290,349} In this model, it was claimed that the conversion comes from the change in active interface. However, this model cannot be used to understand the conversion of F8 and cF8 in asymmetric interface devices, e.g., Au/Sr\textsubscript{2}TiO\textsubscript{4}/Nb:SrTiO\textsubscript{3} cells. Therefore, a second model that combined the oxygen-vacancy-migration and electron-trap-detrap mechanisms has been proposed (see Sec. VC2).\textsuperscript{179} Recently, Lee et al.\textsuperscript{212} suggested a quantitative model for investigating the effect of the nonuniform distribution of oxygen vacancies on the Schottky barrier (see Sec. VC3). They found that oxygen vacancy migration could induce F8 as well as cF8 bipolar switching. In Secs. VC1–VC3, we will review in detail the above-mentioned three mechanisms for coexistence of cF8 and F8 in the same cells.

### 1. Changes in the active interface of a symmetric device

Polarity conversion between F8 and cF8 was observed in a Pt/TiO\textsubscript{x}/Pt cell and explained in terms of the conversion of the active interfaces in a symmetric device.\textsuperscript{290,349} The changes inside the cell are schematically shown in the bars...
of Fig. 43. They assumed that, during RS, there should be three types of layer: the top interface, the bulk, and the bottom interface. A Schottky barrier can be formed at the top and bottom interfaces and either height or width can be controlled by the density of the oxygen vacancies. A metal/semiconductor interface under heavy doping typically exhibits ohmic behavior, whereas low doping results in rectifying (Schottky) behavior. The effective circuits are displayed on the left side of the figure. A lightly doped layer becomes an active layer, which dominates the resistance of the cell, because the largest voltage drop occurs across the Schottky barrier. Therefore, the resistance of the oxide cell is determined by the oxygen-vacancy density at the interfaces. When the oxygen-vacancy density is large at the bottom (top) interface, the top (bottom) interface becomes the active electrode and the cell functions as a forward rectifier, as shown in the top (bottom) right (left) part of Fig. 43.

These different oxygen-vacancy configurations lead to different polarities when oxygen vacancies move in the interfacial layers in response to an applied bias. When the top interface is the active electrode, the application of a negative (positive) bias to the top electrode attracts (repels) oxygen vacancies toward (from) the active electrode, which decreases (increases) the resistance of the cell. Thus, this configuration generates a cF8 hysteresis curve. When the bottom interface is the active electrode, a positive effect occurs; thus the reverse rectifier generates an F8 hysteresis curve. Conversion between the two different polarities can be realized by applying a large electric field. When a sufficiently large negative voltage is applied to the top electrode of the forward rectifier, oxygen vacancies will be attracted to the top interface. Through this process, the cell changes to become a reverse rectifier; the top interface becomes a highly doped layer, whereas the bottom interface becomes lightly doped. The opposite transition, from a reverse to a forward rectifier, is also possible when a large positive voltage is applied to the top electrode.

2. Homogeneous and inhomogeneous switching

Contrasting with the symmetric device explained in Sec. V C 1, it has been observed that asymmetric devices, such as Au/Sr$_2$TiO$_4$/Nb: SrTiO$_3$, also exhibit coexistence of and conversion between F8 and cF8 polarities in a single device$^{179,346,349}$ as shown in Fig. 42(b). In such an asymmetric device, a Schottky barrier is formed at the top electrode, whereas the bottom interface is ohmic; thus, an explanation based on polarity conversion due to changes in the active electrode is not suitable for asymmetric devices. In addition, in the symmetric device$^{290}$ cF8 and F8 share almost the same LRS, while the resistance level of HRS of cF8 is similar to that of LRS of F8 in some asymmetric devices.$^{346,349}$ These facts indicate that a different model is needed for the asymmetric devices.

To explain the coexistence of the two polarities in an Au/Sr$_2$TiO$_4$/Nb: SrTiO$_3$ device, Shibuya et al.$^{346}$ proposed that two different mechanisms are responsible for the two polarities. First, they described the cF8 hysteresis $I$–$V$ curves using the virtual cathode model with a Schottky barrier. A detailed explanation of this model is given in Sec. III C 1 a. Such cF8 $I$–$V$ hysteresis curves were observed over a voltage range of $-4 \ V$ to $+4 \ V$. When the negative voltage exceeded $-4 \ V$, the electrical conductance started to decrease, and the switching polarity was converted to F8. The authors attributed the F8 behavior to trapping of electrons at the interface. They considered that the concentration of positively charged oxygen vacancies at the Au/Sr$_2$TiO$_4$ interface saturates at a bias of approximately $-4 \ V$, and once this is exceeded, electrons are captured in the trap sites, i.e., the oxygen vacancies. This leads to a reduction in the net positive charge in the depleted region, which increases the width of the barrier. Under positive biases, electrons are extracted from the trap states, and the Schottky barrier becomes narrower.

It has subsequently been reported that the observation of both polarities in an asymmetric device may originate from oxygen-vacancy migration. Muenstermann et al.$^{179}$ found the coexistence of both polarities in an asymmetric Pt/Fe: SrTiO$_3$/Nb: SrTiO$_3$ cell, that is similar to those shown in Fig. 42(b). One observed distinction between the two polarities was that they exhibited different electrode-area-scaling behavior. The F8 polarity scaled with the electrode size,
whereas the cF8 polarity had no clear dependence on the size of the electrodes. When the F8 polarity was obtained, the switching area was distributed laterally and uniformly along the interface. With the cF8 polarity, the switching area was localized as a filamentary channel.

To explain the relation between electrode-size scaling and polarity, Muenstermann et al.\textsuperscript{179} proposed the following mechanism. First, consider the case in which the laterally nonuniform filamentary channel is dominant in switching. During the forming process, a non-stoichiometric filamentary channel is generated. The switching mechanism inside the filamentary channel is expected to be the same as that in the conventional virtual cathode model. This leads to cF8 polarity. When a negative voltage is applied to the Pt/Fe-doped SrTiO3 interface, however, oxygen vacancies accumulate at that interface. If this accumulation process is allowed to continue, a region that is depleted in oxygen vacancies forms laterally and uniformly along the interface in the bulk of the oxide cell, because the number of oxygen vacancies in a thin film is limited. This region, which is depleted in oxygen vacancies, is electrically insulating; thus, the cell enters the HRS. A positive voltage applied at the upper interface repels oxygen vacancies from the interface and the LRS state is restored. This generates the F8 polarity.

3. Variation in spatial distribution of oxygen vacancies: The semiconductor with mobile dopant model

As described in Sec. V C 2 , oxygen-vacancy migration is expected to be associated with both F8 and cF8 polarities. To quantitatively describe the role of oxygen vacancies in the resistance of an oxide cell, Lee et al. proposed the SMD model\textsuperscript{212} (see Sec. III C 3 b for more about the SMD model).

The SMD model predicts that the migration of oxygen vacancies near a Schottky interface will result in two different sets of conductance behavior, depending on the spatial distribution of the vacancies. It was assumed that right and left interfaces faced with metals were Schottky and ohmic contacts, respectively. In this asymmetric geometry, the most resistance change comes from the Schottky contact interface, so they could understand how oxygen vacancy distribution near the Schottky interface would determine the polarity of the bipolar switching. Consider the case in which oxygen vacancies are initially distributed far from the Schottky interface, as shown in Fig. 44(a). When a negative (positive) voltage is applied at the Schottky interface, oxygen vacancies are attracted towards (repelled from) the Schottky interface, which results in a decrease (increase) in the width of the Schottky barrier, and the cell enters the LRS (HRS). This leads to cF8 polarity, as shown in Fig. 44(b). This relationship between the motion of the oxygen vacancies and the conductance change corresponds to the conventional explanation in the virtual cathode model.

Now consider the case in which oxygen vacancies accumulate near the interface, as shown in Fig. 44(c). In this case, SMD model calculations predict that attracting (repelling) oxygen vacancies to (from) the interface by applying a negative (positive) voltage results in an increase (decrease) in the width of the Schottky barrier. That is, the SMD model predicts that more (less) doping at the Schottky interface will increase (decrease) the resistance, leading to F8 polarity, as shown in Fig. 44(d). This appears to contradict the conventional semiconductor theory of doping. Note that this result appears to be similar to that of Jeon et al.,\textsuperscript{250} where the Schottky barrier height increased when oxygen vacancies accumulated at the electrode; however, in the SMD model, such penetration through the electrode need not be taken into account. This seemingly contradictory prediction was recently corroborated by experiments in Nb-doped SrTiO3 single-crystal cells,\textsuperscript{197} as explained in Sec. III A 2 a (see Fig. 5(a)).

Based on the above results, the observed polarity conversion\textsuperscript{179,346} is easily understood. When oxygen vacancies are distributed far from a Schottky barrier and the polarity is cF8, if a large negative voltage is applied to the Schottky interface, oxygen vacancies are attracted to the interface and the polarity is converted to F8. If a positive voltage is applied to such a cell, where the oxygen vacancies are distributed primarily near the Schottky interface, oxygen vacancies are repelled from the interface, and the polarity is switched to cF8.

VI. CHALLENGES IN APPLICATIONS

A. Applications of resistive switching to next-generation nonvolatile memories

Following the development of metal oxide semiconductor field-effect transistors, silicon-based devices have dominated the market for over 50 years. One of the driving forces of such a development has been the successful reduction in size of silicon-based devices every year, represented by “Moore’s law.”\textsuperscript{16,17} As the size of such devices approaches that of atoms or molecules in the near future, it is expected that conventional silicon-based devices will face their scaling limits.\textsuperscript{350} Therefore, much effort has been devoted to the development of new scientific and technical concepts and
advanced technologies for next-generation electronic devices.39

Memory is one of the most important building blocks of many modern electronic devices. The basic requirements for next-generation random-access memory (RAM) are nonvolatility, high speed, durability, and low energy consumption. The first part of Table IV presents an overview of the physical mechanisms and operating parameters of commercially available memories (namely, static RAM (SRAM), dynamic RAM (DRAM), and FLASH).38,39,54,351 Both SRAM and DRAM are volatile, and only FLASH can work as a nonvolatile memory. Because such states are stable without metastable states that can be used, the “0” and “1” states of RAM (DRAM), and FLASH).38,39,54,351 Both SRAM and


<table>
<thead>
<tr>
<th>Cell elements</th>
<th>SRAM</th>
<th>DRAM</th>
<th>FLASH</th>
<th>STT-RAM</th>
<th>PCRAM</th>
<th>FeRAM</th>
<th>RRAM (bipolar filament)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage mechanism</td>
<td>Interlocked state of logic gates</td>
<td>Charge on a capacitor</td>
<td>Charge trapped in floating gate or in gate insulator</td>
<td>Magnetization of ferromagnetic layer</td>
<td>Reversibly changing amorphous and crystalline phases</td>
<td>Remnant polarization on a ferroelectric capacitor</td>
<td>Valence change filament formation</td>
</tr>
<tr>
<td>Feature size (F, nm)</td>
<td>45</td>
<td>36</td>
<td>65</td>
<td>45</td>
<td>16</td>
<td>65</td>
<td>45</td>
</tr>
<tr>
<td>Cell area (F&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>140F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>6F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>(12–30)F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>10F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>4F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>20F&lt;sup&gt;2&lt;/sup&gt;</td>
<td>4F&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Read time (ns)</td>
<td>0.2 ns</td>
<td>&lt;10 ns</td>
<td>2 ns</td>
<td>15 ns</td>
<td>0.1 ns</td>
<td>35 ns</td>
<td>12 ns</td>
</tr>
<tr>
<td>Write/erase time (ns)</td>
<td>0.2 ns</td>
<td>&lt;10 ns</td>
<td>2 ns</td>
<td>1 μs/10 μs</td>
<td>1 μs/10 μs</td>
<td>35 ns</td>
<td>100 ns</td>
</tr>
<tr>
<td>Retention time</td>
<td>Volatile</td>
<td>64 ns</td>
<td>4 μs</td>
<td>10 yr</td>
<td>10 yr</td>
<td>&gt;10 yr</td>
<td>&gt;10 yr</td>
</tr>
<tr>
<td>Write cycle (ns)</td>
<td>&gt;1 × 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>&gt;1 × 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>&gt;1 × 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>6 × 10&lt;sup&gt;4&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;3&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;2&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Write operating voltage (V)</td>
<td>1</td>
<td>2.5</td>
<td>2.5</td>
<td>&lt;10</td>
<td>15–20</td>
<td>1.8</td>
<td>3</td>
</tr>
<tr>
<td>Read operating voltage (V)</td>
<td>1</td>
<td>1.8</td>
<td>1.7</td>
<td>4.5</td>
<td>4.5</td>
<td>1.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Array write energy (J/bit)</td>
<td>Unavailable</td>
<td>1 × 10&lt;sup&gt;-13&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-13&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-10&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-11&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-10&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-10&lt;/sup&gt;</td>
</tr>
<tr>
<td>Single cell write energy (J/bit)</td>
<td>5 × 10&lt;sup&gt;-16&lt;/sup&gt;</td>
<td>4 × 10&lt;sup&gt;-15&lt;/sup&gt;</td>
<td>5 × 10&lt;sup&gt;-15&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-10&lt;/sup&gt;</td>
<td>1 × 10&lt;sup&gt;-11&lt;/sup&gt;</td>
<td>2.5 × 10&lt;sup&gt;-12&lt;/sup&gt;</td>
<td>6 × 10&lt;sup&gt;-12&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

As RRAM uses the RS phenomenon as the basic principle of its memory operations, it has several advantages for next-generation memory,38,44,45,51,54,56,352, which are not clearly stated in Table IV. First, a material with RS has two metastable states that can be used, the “0” and “1” states of nonvolatile memory. Because such states are stable without external bias, the operational energy for RRAM can be quite small. Second, in RRAM, the “0” and “1” states can be switched with external electric pulses, rendering its operation is simple and easy. Third, the resistance value of each state can be easily read by applying a very small voltage without disturbing the original state, which allows non-destructive reading. Fourth, the simple capacitor geometry of a RRAM cell without a transistor makes the device highly scalable. Finally, as RS has been observed in numerous insulating materials, it should be easier to find appropriate RS materials that are compatible with the current complementary metal oxide semiconductor technologies than is the case for other memory candidates. Due to these advantages, RRAM has attracted much attention from researchers, and good operating performance for a single cell unit has been already been achieved, as shown in Table IV.

In addition, RRAM devices can be used to overcome the scaling limit of modern silicon-based devices.38,44,45,51,54,56,352 As shown in Table IV, it has already been demonstrated that the scaling limits of RRAM can allow very small feature sizes (F) and a small cell area of 4F<sup>2</sup>/n (n: number of layers in three-dimensional structures). Furthermore, the two-terminal configuration of RRAM devices makes them suitable for a stacked structure.28,29,351 For example, a passive two-terminal stacked structure consisting of units with one diode and one resistor may be considered more advantageous than active configurations that include a transistor (>6–8 F<sup>2</sup>). Such a stacked structure offers significant advantages for increasing memory storage capacity compared with simple two-dimensional layered structures. The stacked RRAM implementation also has the potential to
be combined with other devices to obtain the desired properties of future nonvolatile memory devices.

B. Reducing switching parameter distribution

Although RS phenomena offer many advantages for memory applications, as explained in Sec. VI A, there are several technological difficulties that should be overcome before the actual commercialization of RRAM. One of the major technical obstacles is the wide distribution of switching parameters, including switching voltages.\(^73\)\(^88\) For example, Fig. 45(a) shows 100 \(I–V\) curves in Pt/NiO/Pt cells exhibiting typical unipolar switching.\(^74\) Reset processes occur when voltages of 0.5–1 V are applied and the set voltages fluctuate greatly between 1 and 4.5 V. It should be noted that both set and reset processes in unipolar switching occur under the same voltage polarity. Therefore, these wide distributions make the ranges of set and reset voltages overlap at around 1.0 V, which will cause switching failures. Occasionally, when the set voltage is smaller than the reset voltage, the operation of the reset process can be accompanied by the set process. This operational error is called the “reset failure” phenomenon. When such failures occur, the basic write operations of the RRAM cannot work, possibly leading to severe data-writing/erasing problems in the RRAM. Therefore, reducing variations in switching parameters has been one of the major technological challenges in RS studies.\(^89\)

1. In unipolar switching

Several technical methods have been suggested and examined to overcome the difficulties caused by the wide distribution in unipolar switching voltages: artificial engineering at an electrode–oxide interface; the error-correction algorithm; and the insertion of artificial defects in oxides.

\section*{a. Artificial engineering at an electrode–oxide interface.} Layer-insertion techniques have been used to reduce the variation in switching parameters in unipolar switching devices.\(^73\)\(^74\)\(^79\)\(^82\)\(^85\)\(^354\) Kim \textit{et al.}\(^74\) showed that the wide distribution of switching parameters in Pt/NiO/Pt unipolar switching cells was minimized by inserting a thin IrO\(_2\) layer between the NiO and the Pt electrodes. As shown in Fig. 45(b), the reset failure of the Pt/NiO/Pt disappeared; that is, the set and reset voltages became clearly separated by a gap of \(~0.5\) V. The authors assumed that the CF would be formed via generation of oxygen-related defects by a high electric field, with the possible assistance of the Joule heating effect. Thus, they attributed the reduction in the switching voltage distribution to the IrO\(_2\) layers that stabilized oxygen migration during the formation and rupture of CFs. Jeong \textit{et al.}\(^73\) fabricated failure-free Pt/NiO/Pt cells by forming a very thin Ni–Pt layer at the bottom interface. The authors assumed that the superior oxidation ability of the Ni–Pt electrode compared with that of pure Pt may assist stable RS behavior during the formation and rupture process. Similar experiments were performed on HfO\(_2\) unipolar switching devices by inserting an additional layer at the interface with Al\(_2\)O\(_3\) layer\(^82\) or Ti layer.\(^85\)

The failure-free cells were also fabricated by inserting an artificial structure at electrode–oxide interfaces. Wang \textit{et al.}\(^73\) improved the unipolar switching behavior in a Pt/InGaZnO/Al memory device by creating an arc-shaped bottom electrode.\(^79\) Figures 46(a) and 46(b) show the schematics of cell geometry with and without the arc-shaped bottom electrode, respectively, and the corresponding evolution of CFs. The electric field for the device with an arc-shaped bottom electrode would point to the center of the arc, like a lightning effect, as shown in Fig. 46(a). Then, the CFs were expected to form a simple, radial structure around the arc-shaped bottom electrode. This controlled-CF shape is effective for achieving a narrower resistance distribution and clear switching without multiple transitions in reset and set processes. More complex CF structures are easily formed for the device without an arc-shaped bottom electrode, as shown in Fig. 46(b). This CF results in an unpredictable rupture and a large resistance fluctuation in reset and set processes.

\section*{b. Insertion of artificial defects in oxides.} The fluctuation in switching parameters in unipolar switching can be also reduced significantly by inserting artificial defects in the oxide.\(^76\)\(^78\)\(^80\)\(^83\)\(^86\)\(^355\) The RCB network model revealed that the performance of unipolar switching RRAM can be enhanced dramatically by the presence of defects at specific locations in the percolating network.\(^77\) Figure 47(a) shows the distribution of the forming, reset, and set biases in which the pristine on-state circuit breakers were chosen randomly.
in the RCB network. The three distributions exhibited large variations, which explain the commonly observed fluctuation in switching voltages. Similar simulations were repeated by inserting artificial defects, in this case by setting four-aligned on-state circuit breakers near the top electrode in the pristine state. As shown in Fig. 47(b), the variation in the forming, reset, and set processes was significantly reduced. This fluctuation reduction originated from reducing the variation in the shape and location of the CFs.

This idea has actually been demonstrated experimentally. As shown in the inset of Fig. 47(d), Yoon et al. embedded Ru nanodots near the top Pt electrodes in Pt/TiO₂/Pt cells using atomic layer deposition.⁷⁸ Figures 47(c), 47(d), and 47(e) show the cumulative probabilities of the set voltage, reset power, and reset voltage, respectively. When the areal density of Ru nanodots increased (i.e., the number of deposition cycles increased), the switching parameters become more uniform.

c. Error-correction algorithm. To overcome the failure problem, an error-correction algorithm could be used for RRAM.⁸⁴ In fact, similar methods have been widely used in conventional memory devices, including FLASH memory and DRAM.⁵⁴ Lee et al.⁷⁵ demonstrated that the "reset fail" phenomenon can be avoided by using a software error-correction algorithm instead of the additional layer in the RRAM structure. Figure 48(a) shows their measured plots of \( V_{\text{reset}} \) and \( V_{\text{set}} \) according to the switching cycle number, measured from unipolar NiO cells, where \( V_{\text{reset}} \) and \( V_{\text{set}} \) have a very wide distribution. The voltage overlap during successive RS operations is clearly seen, as indicated by circles. The largest value of \( V_{\text{reset}} \) and the smallest value of \( V_{\text{set}} \) were 1.39 and 1.12 V, respectively, indicating that \( V_{\text{reset}} \) and \( V_{\text{set}} \) sometimes overlapped with each other near 1 V. Such overlap should induce switching failure.

Lee et al.⁷⁵ proposed that the reset current \( I_{\text{reset}} \) value for the next reset process could be predicted using the scaling relationship between \( I_{\text{reset}} \) and the third harmonic signal \( B_{3f} \), which is outlined in Table III and Sec. IV C 1 b. In terms of scaling relation, they used \( B_0 \), which is the dc limit for \( B_{3f} \). Here, \( B_0 \) can be estimated from the fitting curve \( V/I = R_o + B_0 I^2 \) using the experimentally measured \( I-V \) curve in the LRS, where \( R_o \) is the resistance value of the LRS at the limit of zero-external voltage. Namely, each \( B_0 \) can be obtained from each LRS \( I-V \) curve. Without actually performing the next reset process, the forthcoming \( I_{\text{reset}} \) value, \( I_{\text{reset}}(\text{Pre}) \), can be predicted by using the scaling relationship, \( I_{\text{reset}}(\text{Pre}) = C B_o^{-\eta} \), where \( C \) and \( \eta \) are experimentally determined parameters. The value of \( V_{\text{reset}} \) for the next reset process, \( V_{\text{reset}}(\text{Pre}) \), can also be predicted from \( V_{\text{reset}}(\text{Pre}) = I_{\text{reset}}(\text{Pre}) R_o = B_0 I_{\text{reset}}(\text{Pre})^2 \).

Using the estimated \( V_{\text{reset}} \) (Pre), Lee et al.⁷⁵ developed an error-correction scheme for RRAM by controlling the maximum value of the allowed \( V_{\text{reset}} \) distribution. They defined \( V^{th}_{\text{reset}} \) as a preassigned threshold voltage for the maximum value of \( V_{\text{reset}} \). For a given LRS cell, when \( V_{\text{reset}} \) (Pre) > \( V^{th}_{\text{reset}} \), the LRS should be discarded because it is highly probable that the real \( V_{\text{reset}} \) will be larger than \( V^{th}_{\text{reset}} \). Therefore, this LRS has to be erased by the subsequent set process, which makes the device enter a new LRS. Then, the above process should be repeated until \( V_{\text{reset}} \) (Pre) < \( V^{th}_{\text{reset}} \). Through this algorithm, it is expected that the maximum value of \( V_{\text{reset}} \) will be smaller than \( V^{th}_{\text{reset}} \). This algorithm was actually applied to the PtNiO/Pt cells exhibiting the overlap behavior, as shown in Fig. 48(a). The results after applying the error-correction scheme are shown in Fig. 48(b). They clearly demonstrate that \( V_{\text{reset}} \) is always smaller than \( V^{th}_{\text{reset}} \) and that the error-correction method can lead to a clear voltage window between \( V_{\text{set}} \) and \( V_{\text{reset}} \). Thus, this error-correction scheme guarantees that the reset failure due
to voltage fluctuation in unipolar switching can be overcome.

2. In bipolar switching

The aforementioned techniques used for unipolar switching cells can be applied to reduce the fluctuations in switching parameters in bipolar switching. To improve the switching parameter distributions, Lee et al. and Sadaf et al. fabricated ZrO$_x$/HfO$_x$ and WO$_x$/NbO$_x$ bilayer structures, respectively, which is essentially the same technique as the layer insertion described earlier in Sec. VI B 1 with regard to unipolar switching. In these devices, the improvement in the fluctuations was attributed to the formation of a single CF in one of the two layers, which does not disappear even in the HRS. This CF can effectively act as an artificial defect or a lightning rod, which stabilizes the bipolar switching operations in the other oxide layer. Experiments involving insertion of artificial defects were also performed in bipolar Ag/ZrO$_2$/Pt devices. By inserting Cu-nanocrystals near the bottom interfaces, bipolar switching properties developed excellent uniformity. This enhancement was interpreted as the Cu-nanocrystals acting as seeds for nucleation and growth of conducting channels.

To reduce the distribution of switching parameters in bipolar switching, Lee et al. proposed a new architecture using vertically aligned heterointerfaces. It has been well established that, when more than one immiscible oxide is deposited in a simple film form, a very simple and self-assembled structure, called a nanoscaffold film, can be fabricated. Figure 49(a) shows a scanning TEM image of a nanoscaffold film containing Sm$_2$O$_3$ nanocolumns. Atomic arrangement at the vertical interface can be crystallographically modelled. Figure 49(b) shows the cross-sectional TEM image, which depicts the atomic arrangement at the vertical heterointerfaces of nanocolumns and matrices. Due to the structural incompatibility of Sm$_2$O$_3$ and SrTiO$_3$, there should be many oxygen deficiencies at the vertical heterointerfaces.
the oxygen migration highway. Bipolar switching was observed without the forming process at room temperature. As shown in Fig. 49(c), resistance variations exceeded two orders of magnitude, with extreme uniformity. Using numerical simulations, the bipolar switching in nanoscaffold films was explained by the Joule heating-accelerated drift of oxygen vacancies, localized at vertical heterointerfaces. Therefore, local ionic conduction along vertical heterointerfaces can reduce the wide fluctuation in switching parameters.

C. Multilevel resistive switching

Throughout this review, we have explained that there are two resistance states, LRS and HRS, in RS phenomena, which can be used as binary states for memory devices. In fact, most resistive memory cells can exhibit more than two resistance states between LRS and HRS in a single cell, which can be applied to multibit memories. This can increase the storage capacity, that is, the information density, can increase for a given number of RS cells.

There are two typical ways to obtain multiple resistance states in RS phenomena. The first is by varying the defect concentration during switching, which is usually used in bipolar switching. Because the resistance of a cell in bipolar switching is determined by the concentration of oxygen vacancies or electron-trap sites, as explained in Sec. III C, if we can vary their concentrations, it is possible to have multiple resistance states. This can be accomplished by applying different external biases to the bipolar cells; a larger bias can induce a larger concentration change. Figure 50(a) shows such multilevel RS of an epitaxial Pt/(Ba,Sr)TiO$_3$/SrRuO$_3$ bipolar switching cell. When biases of 2, 3, and 5 V were applied successively, stable resistance states of 40, 60, and 85 kΩ, respectively, were measured. The three different resistance states can form bits as they are clearly distinguishable.

The second way to obtain multiple resistance states is by varying the interconnection inside the CFs. This is usually used in unipolar switching. Because the resistance of the cell is strongly dependent on the interconnection between the conducting regions inside the CFs, if we gradually vary the interconnections, we obtain multiple resistance values. This can be achieved by controlling the compliance current or the external voltages during the switching operation. Seo et al. reported multilevel resistance states in unipolar switching NiO films. By changing the compliance current from 1 to 20 mA, they found that the resistance of the LRS varied linearly with the compliance current, which originates from the variation in CF structure. Similarly, by carefully controlling the bias around the reset voltage, Chae et al. found that the resistance of unipolar switching TiO$_2$ films could vary over a range between those of the LRS and the HRS. Using C-AFM, they revealed that multilevel RS was accompanied by changes in the area and conductance of the conducting regions at the top surface, as shown in the upper figures of Fig. 50(b). This systematic behavior could be successfully confirmed by simulations using the RCB network model, as shown in the lower figures of Fig. 50(b): the red, orange, and green-colored regions show the area of the conducting regions at the top surface. These simulations indicate that different resistance states can be obtained from partial rupture of CFs.

D. The sneak-path problem in crossbar arrays

As device features are pushed towards the deep sub-100-nm regime, the conventional scaling methods of the semiconductor industry face increasing technological and fundamental challenges. For example, fluctuations in device size may result in a large spread in device characteristics at the nanoscale, affecting key parameters such as the operational voltages and on/off currents. Therefore, crossbar architecture, shown in Fig. 51, is a new paradigm for RAM technology; it has been lauded for its simplicity, robust switching, promising performance metrics, and its potential for terabit-scale density. A crossbar structure is composed of a set of parallel bottom electrodes, called bit-lines, and perpendicular top electrodes, called word-lines. The word- and bit-lines sandwich memory devices. When $+V_{ \text{cell}}/2$ and $-V_{ \text{cell}}/2$ are applied to one of the word-lines and one of
When two Ta$_2$O$_5$ layers are connected in an anti-serial configuration, the current is reduced near the origin of the I–V curve. For oxide-resistive memories, the current is applied between bit-line A2 and word-line B3. The current is expected to flow from A2 → cell #1 → B3, as shown by the solid line in the figure, to read the data from cell #1. However, due to surrounding low-resistance cells, the current flows through them, as shown by the dashed line, leading to cell #1 being incorrectly read as in a low-resistance state. In principle, the sneak-path problem can be overcome by incorporating active elements such as transistors with RS cells; under such conditions, all bypass currents can be blocked by controlling the transistors. However, the introduction of these additional active devices and related circuits significantly reduces integration density.

It is highly desirable to find the simplest structure that solves both the sneak-path and the memory density problems. To this end, several device structures combining two different switching elements into one have been proposed. Linn et al. addressed the sneak-path problem by connecting two resistive memory cells in an anti-serial configuration. The I–V curves of bipolar switching are asymmetric at the origin for positive and negative voltages, as shown in the left portion of Fig. 52(a). When two resistive memories are connected in an anti-serial configuration, the current is reduced near the origin of the I–V curve, as shown in the right portion of Fig. 52(a). Anti-serial connection was realized for oxide-resistive memories using double-layered Ta$_2$O$_5$–TiO$_2$–x cells. Each of the double-layered Ta$_2$O$_5$–TiO$_2$–x cells showed asymmetric bipolar switching I–V curves, similar to the left portion of Fig. 52(a). When two Ta$_2$O$_5$–TiO$_2$–x cells are anti-serially connected, the overall I–V curves should be similar to the right side of Fig. 52(a). That is, the current level near V = 0 V is suppressed.

Switching elements, so-called selectors, can also block leakage currents to facilitate the correct reading of the desired cell. Figure 52(b) shows a schematic illustrating the RS layer connected to a switch in one cell. This switch can suppress the current when the applied voltage is less than the switching voltage. For unipolar RRAM, a VO$_2$ switch can be used to overcome the sneak-path problem in unipolar RRAM; Lee et al. demonstrated that the sneak-path current in a NiO memory cell was suppressed at low voltages by connecting memory cells to the VO$_2$ switch. A similar approach has been taken with regard to bipolar RRAM. Chang et al. realized this concept by connecting TiO$_2$–resistive memory to a VO$_2$ switch. Unlike the anti-serial connection method, the intermediate electrode is not necessary in this method. This approach has been applied to other threshold-switching and resistive-switching devices, including NbO$_x$–WO$_x$ cells and WO$_x$–VO$_2$ cells.

VII. CONCLUSION

In this article, we reviewed the basic mechanisms of RS phenomena observed in oxide materials. Instead of covering all the detailed mechanisms that are applicable to specific materials, we tried to provide general insights into RS mechanisms by focusing on the roles of common defects in oxides.
Oxygen vacancies were our major concern at the microscopic level, because they are intrinsic defects in oxides and, in general, strongly affect the transport behaviors of oxides. Therefore, based on various experimental and theoretical studies, we first reviewed how local electric resistance can be changed by the microscopic change in the oxygen vacancy configuration. The important microscopic mechanisms were: (i) the formation and rupture of CFs by aggregation and dissociation of oxygen vacancies; (ii) the modulation of the Schottky barrier driving forces of oxygen vacancy migration. (iii) the trap-and-detrap of electrons in oxygen vacancy sites. We also discussed the roles of the electric field and Joule heating as the driving forces of oxygen vacancy migration.

However, in this review, we have demonstrated that such microscopic studies are limited in their ability to provide a sufficient understanding of the observed statistical properties of RS behaviors. Note that defects, such as oxygen vacancies, are not arranged periodically as in a solid, but are instead distributed nonuniformly. Thus, the oxygen vacancy configuration in a given sample can differ from that in another, although the experimental conditions are the same. Additionally, when oxygen vacancies are clustered in a local region, they form a CF. The difference in the oxygen vacancy (i.e., CF) configuration is the major reason for the wide distribution in the fluctuations of the switching parameters. Therefore, it is of critical importance that we understand the collective behaviors of oxygen vacancies and the statistical properties of the diverse patterns of the CF configurations.

For this reason, we used a statistical physics to provide an overview of the RS research. Specifically, percolation theory, which enables us to investigate RS phenomena at the mesoscopic level, is among the important approaches in this domain. In contrast to microscopic approaches, material-independent modeling is possible at the mesoscopic level, regardless of microscopic detailed mechanisms. Therefore, various statistical features observed in RS phenomena and some switching-type conversion phenomena were well explained quantitatively and qualitatively by the percolation models. In addition, we have shown that the statistical physics approaches help us to discover that there are universal scaling laws behind seemingly random noise-like fluctuations in RS phenomena that originate from the material-independent structure of the CFs.

These studies also indicate the importance of oxygen vacancy configuration engineering when viewed from statistical physics perspectives. As reviewed in Section VI, the wide fluctuations in switching parameters can be reduced by controlling the CF shapes. Because the reduction of the large fluctuations is the major issue for the practical use of RS phenomena as memory devices, we hope the statistical physics approaches will provide a fundamental theoretical basis for RRAM technology in the future. Considerable research effort based on these approaches is needed to further explore the potential application of RS phenomena and to exploit their operational limits.

ACKNOWLEDGMENTS

We are very grateful to Professor Seung Chul Chae and Professor Seo Hyoung Chang for their careful reading of our manuscript and helpful comments. This work was supported by the Institute for Basic Science (Grant No. IBS-R009-D1) and the National Research Foundation of Korea (Grant No. NRF-2011-35B-C00014 to J.S.L.). S.L. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division.

2D. P. Tseng, Ferroelectrics 471, 23 (2014).
12H. Bieckman, Vacuum 26, 513 (1976).