REVIEW ARTICLE

Access devices for 3D crosspoint memory[a)]

Geoffrey W. Burr,[b)] Rohit S. Shenoy,[c)] Kumar Virwani, Pritish Narayanan, Alvaro Padilla,[d)] and Bülent Kurdi
IBM Research—Almaden, 650 Harry Road, San Jose, California 95120

Hyunsang Hwang
Pohang University of Science and Technology (POSTECH), Materials Science and Engineering, 77 Cheongam-ro, Nam-gu, Pohang, Gyeongbuk 790-784, South Korea

(Received 17 June 2014; accepted 20 June 2014; published 24 July 2014)

The emergence of new nonvolatile memory (NVM) technologies—such as phase change memory, resistive, and spin-torque-transfer magnetic RAM—has been motivated by exciting applications such as storage class memory, embedded nonvolatile memory, enhanced solid-state disks, and neuromorphic computing. Many of these applications call for such NVM devices to be packed densely in vast "crosspoint" arrays offering many gigabytes if not terabytes of solid-state storage. In such arrays, *access* to any small subset of the array for accurate reading or low-power writing requires a strong nonlinearity in the IV characteristics, so that the currents passing through the *selected* devices greatly exceed the residual leakage through the *nonselected* devices. This nonlinearity can either be included explicitly, by adding a discrete *access device* at each crosspoint, or implicitly with an NVM device which also exhibits a highly nonlinear IV characteristic. This article reviews progress made toward implementing such access device functionality, focusing on the need to stack such crosspoint arrays vertically above the surface of a silicon wafer for increased effective areal density. The authors start with a brief overview of circuit-level considerations for crosspoint memory arrays, and discuss the role of the access device in minimizing leakage through the many nonselected cells, while delivering the right voltages and currents to the selected cell. The authors then summarize the criteria that an access device must fulfill in order to enable crosspoint memory. The authors review current research on various discrete access device options, ranging from conventional silicon-based semiconductor devices, to oxide semiconductors, threshold switch devices, oxide tunnel barriers, and devices based on mixed-ionic-electronic-conduction. Finally, the authors discuss various approaches for self-selected nonvolatile memories based on Resistive RAM. © *2014 American Vacuum Society*.
[http://dx.doi.org/10.1116/1.4889999]

I. INTRODUCTION

A crosspoint array consists of a lower and an upper plane of closely spaced parallel wires, running at right angles to each other, and the dense, Cartesian array of interconnections at each crossover point between the wires.[1] If both the wires and the spaces between them have a width of F, then the area per connection is $4F^2$. This "crosspoint" concept has been investigated for memory applications for more than 60 yrs.[2]

Such a memory ought to be capable of extremely high densities, with an effective footprint of $4F^2/L$ per memory element for L crosspoint arrays stacked in 3D. Despite this, crosspoint memory has not yet met with widespread commercial success.[3] One reason is that a crosspoint memory application requires a very diverse and very large set of critical characteristics. At each interconnection, there should be both a memory device that stores data ("remembers the strength of the interconnection") and a strong nonlinearity. This nonlinearity is needed to allow external access to just a few of these memory devices without disturbing any of the other memory devices or inducing wasteful leakage through the thousands or millions of other crosspoint intersections.

A crosspoint memory element should be a two-terminal device that can be reliably, repeatedly, and readily switched between at least two resistance states, preferably with a large resistance contrast. This switching operation must not require excessive power—otherwise, only a few bits can be written in parallel and, as a result, the effective write bandwidth will be extremely low. Such a device should also be capable of surviving through many millions, if not billions or more, of switching cycles (write endurance).

[a)]Some portions of this review article will appear in G. W. Burr, R. S. Shenoy, and H. Hwang, "Select device concepts for crossbar arrays," in *Resistive Switching—From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*, edited by D. Ielmini and R. Waser (in press), Chap. 23. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.
[b)]Electronic mail: gwburr@us.ibm.com
[c)]Now with: Intel, Santa Clara, CA 95054.
[d)]Now with: SanDisk, Milpitas, CA 95035.

Each read operation, at lower voltages and currents than the switching operation, should supply current sufficient for distinguishing the various states. Yet even a very large number of successive read operations must not induce an undesired switching event (known as a "read disturb"). The resistance states should be nonvolatile over the normal operating range of the system (typically up to 85 °C for nonmilitary applications) over a long lifetime (often 10 yrs is specified). Error rates should be both predictable and low enough throughout the device lifetime for correction by error-correction coding (ECC) with a reasonable amount of redundancy overhead.

In the past 10–15 yrs, memory devices capable of these kinds of specifications have moved closer to becoming a reality. Research begun in the late 1990s, originally motivated by the desire to have a "backup device" in case either NAND or NOR Flash had trouble scaling to smaller dimensions, has produced a number of promising two-terminal nonvolatile memory (NVM) devices. These include ferroelectric RAM (FeRAM),[4] phase-change memory (PCM),[5] spin-torque-transfer magnetic RAM (STT-MRAM),[6] and resistance RAM (RRAM).[7] Even though NAND Flash succeeded at scaling and, as of early 2014, has not required "replacement," there has been a general realization that many of these two-terminal NVM devices might offer better performance through much lower latencies, direct byte-level addressability, and much higher program-erase endurance than NAND Flash. This opportunity is referred to collectively as *Storage Class Memory*.[8,9]

Ideally, we would have a vast matrix of densely packed devices, within which we could write and read any small subset of the memory devices at will, while all other devices would remain completely unperturbed, dissipating zero additional power beyond that required for writing or reading. While this ideal situation is not really attainable, one can come close by introducing a strongly nonlinear IV characteristic—an *access device*—together with each NVM element at every crosspoint. One way to introduce this nonlinearity is to integrate a second, discrete two-terminal device—such as a diode, switch, or other similar element—together in series with each state-holding element. This approach has the advantage that the memory element and the access device can each be optimized separately, and then integrated together in the final semiconductor processing scheme. However, this approach tends to require additional processing steps, often including costly lithography steps.

A second approach for introducing the necessary nonlinearity is to simply add it to the list of necessary criteria that the prospective state-holding memory element must deliver. This has a distinct processing advantage—once successfully developed, there is only one device to integrate; it has the disadvantage that the list of criteria for state-holding devices is quite daunting already, without adding a strong nonlinearity of many orders of magnitude—at exactly the "right" voltage—to the list.

In this article, we begin with a brief overview of circuit-level considerations for crosspoint memory arrays and discuss the role of the access device in minimizing leakage through all the various nonselected cells, while delivering the right voltages and currents to the selected cell. We then summarize the list of criteria that an access device must fulfill in order to enable crosspoint memory. We review current research on various discrete access device options, ranging from conventional silicon-based semiconductor devices, to oxide semiconductors, threshold switch devices, oxide tunnel barriers, and devices based on mixed-ionic-electronic-conduction (MIEC). Various approaches for self-selected RRAM are discussed, followed by conclusions.

II. CROSSPOINT ARRAY CONSIDERATIONS

The task of operating a crosspoint array is fairly simple: we need to apply a set of voltages at the edge of the array such that the desired operations (reads and writes) take place at the desired *selected cells* (where a "cell" refers to the combination of an access device and memory element at any given crosspoint); yet, all nonselected cells remain unperturbed and the overall power dissipation remains manageable. Over the past few years, a large number of crosspoint analyses have been performed,[10–23] which all show that it is nearly impossible to implement a crosspoint array of any practical size without a strong nonlinearity at each crosspoint—we will not repeat such analyses here.

The overall memory chip, storing anywhere from megabytes to terabytes of memory, will be composed of numerous crosspoint subarrays, which might each be on the order of 1000×1000 wires in size. We want each of these sub-arrays to be as large as possible, such that much of the peripheral circuitry (including address decoders, data buffers, sense amps, and control circuitry) can be placed underneath the arrays, thereby reducing their silicon "foot–print." The larger the subarray, the larger is the *area efficiency*, or the fraction of silicon area associated with memory bits, and higher area efficiency typically implies lower cost-per-bit.

A. Problems associated with large subarrays

There are several drawbacks to using large subarrays, which can partially counteract or even overwhelm the area-efficiency benefits. The longer the wire, the larger the worst-case resistance drop and the larger the capacitance. These problems are exacerbated at advanced technology nodes, as wires get thinner, increase in line resistance, and are spaced more closely together (which increases coupling capacitance[24]). The worst-case access is a write operation at the "far" corner of the array, since the currents are largest during writes, and since both extra power and voltage drop are consumed in the wiring leading out to the cell, and in the orthogonal wire leading back from the cell to the other edge of the array. The need to boost the voltage at the edge of the array then runs the risk of perturbing the state of memory elements in nonselected cells that sit nearest to those edges. Depending on how the vias (between the memory wiring layer and the underlying drive circuitry) are organized, the exact location of the worst-case device can vary. But the general trend remains unchanged: problems associated with line resistance get worse as the subarrays get larger.

Array size and switching current of the memory element must therefore be consistent with ensuring a manageable IR drop on selected lines. For example, a 30 μA current in a 1000×1000 subarray at the 32 nm technology node corresponds to an IR drop of $\sim$66 mV.[24] While this may not seem very high, access device turn-on tends to be highly nonlinear and even small increases in the total voltage that must be applied at the array periphery can significantly increase leakage currents. Furthermore, high write bandwidth may necessitate a parallel write operation across multiple selected cells along a wire. The effect of IR drop can be even more severe in such scenarios.

Write times and bandwidth can also be affected by increased line resistance and capacitance. The large capacitance of long bitlines is a well-known issue in memory design. In fact, it can easily be the wire parasitics that determine the effective memory speed and not the physical memory element itself. These effects are compounded by increases in coupling capacitance as spacing is reduced. Also, larger arrays make distinguishing low and high resistance states (LRS, HRS) during read operation much more difficult for memory elements with low resistance contrast (such as STT-MRAM), or with absolute resistance values comparable to the aggregate line resistance.

An added complication for RRAM is the critical importance of *compliance* currents. Often the RRAM "SET" operation involves triggering of a positive feedback effect (due to the interplay between local temperature, local field, and the underlying electrochemistry), leading to the creation or completion of a filament (of either oxygen vacancies or metal cations). Frequently this filament formation process is terminated not by the device itself, but by the response of the surrounding circuitry. Failure to terminate the filament formation at a particular targeted current can lead to excessively thick filaments, which then require significant switching power to sever (or remove) during the subsequent "RESET" operation.

When large bitline capacitances are present, a transistor or other circuit at the edge of the array can no longer instantaneously limit the current passing through a cell within the array. What is worse, the degree to which the compliance control will be affected can vary across the array as a function of the distance between the circuit that is enforcing compliance and the memory device that is undergoing programming. This is yet another problem that makes it difficult to implement large subarrays.

B. Considerations during NVM-write

On top of difficulties arising from wire parasitics and ensuring that a significant fraction of the externally applied voltage drop appears across the selected cells, we must consider the aggregate power consumption through all the other cells in the activated subarray. These metrics depend directly on the properties of the access device—specifically on its ability to deliver high currents through selected cells while at the same time limiting leakage through possibly millions of nonselected cells. As we will see, however, the bias voltages at which the access device must perform these roles depend strongly on the NVM element with which it is paired.

Typically, a large potential difference is applied across selected cells by driving the corresponding wires to high and low voltages. Voltages through all other wires in the array are set at some intermediate values to minimize leakage and to prevent undesired disturbances to previously stored data. The choice of this set of voltages is known as the biasing scheme. Since it is only voltage differences that matter, the absolute values of all voltages can be chosen to simplify the peripheral driving and supply circuitry.

As shown in Fig. 1, cells in the array which are not fully selected fall into three classes: the *half-selected* cells along the same row as a selected cell, the *half-selected* cells along the same column as a selected cell, and *un-selected* cells that share neither a row nor a column with any selected cell. We

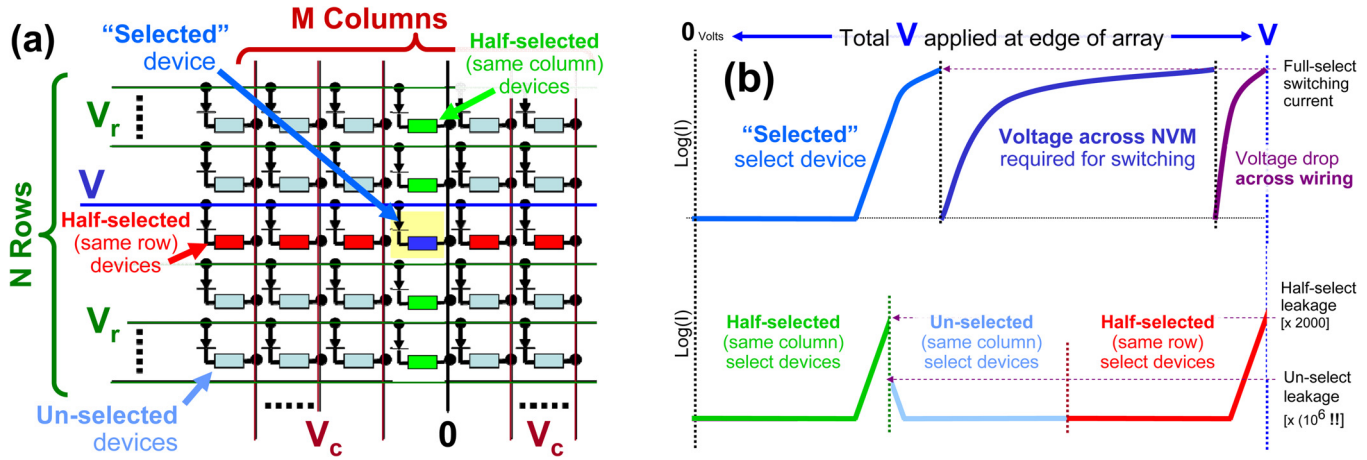


FIG. 1. (Color online) (a) In a crosspoint array of NVM devices, voltage applied at the edge of the array, to select one or more NVM/access device pairs, also affects three other sets of NVM + access devices: those in the same row, in the same column (both *half-selected*), and all others (*unselected*). (b) The total voltage drop across any selected NVM device must be sure to switch the device, despite the additional voltage drop in its own access device and in the wiring. However, that same applied voltage must not lead to excessive leakage in either the half-selected or unselected devices. In a 1000×1000 array, for example, there will be $\sim$2000 half-selected devices and nearly 1×10^6 unselected devices (Ref. 25).

refer to cells as *half-selected* not so much because they see half of the voltage drop across the selected device, although that situation can occur. Instead we can consider that these devices will always be touching one of the two (i.e., half of the) strongly energized wires associated with a selected cell.

We first address the scenario of writing the worst-case device, since this is the largest applied voltage differential that will ever be used. One biasing scheme that works well for this scenario is referred to as the "V/3" scheme.[10–23] In its simplest form, the applied voltage is split into three parts, one for each of the three classes of nonselected devices. Both sets of half-selected devices see the same voltage polarity as the selected device, while the unselected device experiences the opposite voltage polarity. As shown in Fig. 1, the sum of these three voltage drops is roughly equal to the effective voltage drop across the selected device. Although this simplistic analysis is highly useful, the presence of significant line resistance complicates the situation considerably, since the wire and local device voltages then vary continuously along the selected rows and columns.

While the name "V/3" would suggest division into three equal voltages,[26] this is rarely the optimal choice. When we are selecting just one device out of an $N \times M$ array, there are many more $[(M-1)\times(N-1)]$ unselected devices than there are half-selected $[(M-1)+(N-1)]$ devices. In the case of a bipolar access device with symmetric characteristics, it makes sense to reduce the voltage across the many unselected devices until total aggregate leakage is minimized. This scenario is illustrated in Fig. 1. In a scenario where the access device and the NVM device are unipolar (using only one polarity for all switching operations), then it might be preferable to significantly increase the opposite polarity voltage across the unselected devices. This of course assumes that one can depend on the unipolar access device to strongly suppress leakage current even at large "opposite" polarity voltages.

During write operations, it is critical to deliver the necessary power and current to the selected cell while avoiding two undesirable outcomes: dissipating too much aggregate power and risking even the remote possibility that any of the nonselected cells might accidentally switch. The probability of such a *write disturb* during any one write event needs to be extremely low—data stored early in the lifetime of the array must persist even if there are a very large number of writes to all other portions of the array. Any excess power dissipated during write will necessitate a reduction in the write parallelism and thus will reduce the effective write bandwidth. Obviously, as the size of the subarray gets larger, the number of potentially leaky devices increases. In addition, the energy required to drive the wires to the desired voltages and to activate the subarray increases as $(1/2)CV^2$, thus favoring low wire capacitance and low voltage operation. Note also in Fig. 1(b) that the switching voltage of the NVM figures prominently in the calculation of leakage current.[25] As a result, it is difficult if not impossible to accurately quantify the prospects of an access device in isolation, without knowing what NVM with which it is to be paired.[25] For this very reason, there is no quantitative comparison table included with this review article.

A final consideration for write operations is that the narrow wires associated with advanced technology nodes are subject to failure due to electromigration. Passing $\sim 10\,\mu$A current through an F $\sim$ 10 nm width wire with 2.1× aspect-ratio (suitable for 20 nm pitch arrays in a future technology node) would translate to a current density of $\sim$4.5 MA/cm^2—higher than the expected limit on current density imposed by electromigration concerns (3.26 MA/cm^2).[24] Fortunately, this limit has been specified for microprocessor logic circuitry, where the frequency of large switching currents is much higher than the frequency of write operations in an NVM crosspoint array. Thus NVM switching currents of the order of tens of microamps, consistent with some experimentally demonstrated RRAM write currents, may in fact still be feasible.

C. Considerations during NVM-read

During read operations, both the currents and the associated device voltages are usually significantly lower. Since read currents are typically at least 10× smaller than write currents, line resistance can frequently be ignored during read operations. Many authors have studied the use of a "V/2" scheme, where the voltage applied to the wires leading to *unselected* devices is either brought to zero or alternatively, these lines are allowed to float.[14] There are different tradeoffs associated with these choices. Driving these lines to the same voltage has the advantage that the leakage within the array is sure to be low; however, the act of driving these wires consumes some amount of power in the peripheral circuitry.

During write, we were concerned with perturbing stored data and with wasted power; during read, we are additionally concerned with being able to accurately sense the state of individual devices. Without strong nonlinearity at each node, there are many potential "sneak paths" from the higher potential selected wire to the lower potential selected wire. In fact, every path that leads through a series combination of three cells in the low-resistance state—one each from the pool of half-selected/same-row, unselected, and then half-selected/same-column devices—can contribute, at the read-out node, significant signal that is not representative of the state of the selected device.[10–23] Given the large number of possible connections for any nontrivial pattern of stored data, the *sensing margin*—the detectable difference at the edge of the array between the high and low resistance states—collapses in the absence of a strong nonlinearity, even for a fairly modest size array. The presence of such a nonlinearity cuts off the sneak paths, either by suppressing the current through all three legs or in some cases, simply by blocking the current through the reverse-biased *unselected* devices, making it possible to have reasonable sense margin for large arrays.

D. Failed access devices

One consideration that was a frequent topic in the molecular circuits literature,[27] but which is rarely considered in the NVM literature is the presence of failed access devices within the array. Access devices that fail as opens are not the

issue here—although the associated stored bit is inaccessible and ECC must be present to correct the erroneous data, there is no adverse effect on the rest of the array. However, access devices that fail as a low resistance, or "short," can create havoc in a crosspoint array. Since very little voltage falls across the failed access device, the associated NVM device now has a much larger voltage drop than intended, and thus passes a large amount of current. Not only is the state of this device perturbed for almost any access to the array, but this crosspoint intersection is now leaking an unexpectedly large amount of current. This is one factor that would seem to favor driving every wire to a known voltage, rather than depending on wires to "float" to a desired voltage, since shorted devices will shift the potential of such floated lines and lead to unintended device activity all over the array.

However, even when the biasing scheme does call for driving each line to a known voltage, "shorted" access devices can impact both write and read operations. Such failed access devices draw high current, causing writes to draw much more power than expected. Since these high currents can degrade the voltage drop across the selected crosspoint cell(s), write operations that would otherwise have succeeded could potentially fail. By acting as strong "sneak paths," failed access devices also mask read currents from all devices that share the same sense-amplifier during a read operation, degrading read margin. If the number of failed devices is very low, then known techniques for remapping failed static or dynamic random access memory lines can be used to simply redirect that address to a redundant line included for just such purposes. While this removes the possibility that the failed access device will be addressed in either the *selected* or *half-selected* roles, higher leakage will still be generated because every access to the subarray still places the failed device into the *unselected* state.

Another approach, demonstrated for two-terminal MIEC devices in series with phase-change memory (PCM),[28] is to place the state-bearing device in an extremely high resistance state. This is possible for PCM because negative polarity pulses that are ramped down slowly are known to lead to phase-separation of the phase-change material,[29,30] producing a very high resistance state. Although this requires a voltage similar to very large writing voltages, this corrective action can be taken without damaging neighboring devices since the failed access device is, by definition, not consuming any of the applied voltage. The PCM device, once in this ultrahigh resistance state, then protects the array from the leakage that would otherwise be induced by the shorted access device. Unfortunately, for other NVMs such as RRAM, it is not yet clear if devices can be placed in such a high-resistance state with voltage sequences that can be applied within the array without damaging neighboring healthy device pairs.

III. TARGET SPECIFICATIONS FOR ACCESS DEVICES

An ideal access device has a number of critical characteristics for application in high density resistive crosspoint memory arrays.

In each selected cell, the access device should be capable of supplying high currents as needed to program and erase the memory element that is in series with it. Delivering such high currents through the low cross-sectional area corresponding to devices of $\sim 1F^2$ implies that the access device should support a **High ON-state current density**, on the order of several MA/cm^2. For example, passing ~ 10 μA current (consistent with experimentally demonstrated RRAM write currents) through an F $\sim$ 10 nm diameter access device (suitable for 20 nm pitch arrays) translates to a current density of ~ 13 MA/cm^2.

Since the unselected cells vastly outnumber the selected cells in typical crosspoint memory arrays, the access device's **OFF-state leakage current needs to be as low as possible** so as to prevent the aggregate leakage through all the unselected cells from dominating the overall system power budget. Access device OFF-state leakage currents that are several orders of magnitude lower than the ON-state current allow the use of large arrays which are desirable to increase overall area efficiency. For instance, assuming that a single bit is being written in a 1000×1000 array, an access device ON/OFF current ratio of $\sim 10^6$ is needed in order to keep the aggregate unselected cell leakage current to the same order of magnitude as the write current. As the array size increases, the corresponding maximum tolerable OFF-state leakage per device must commensurately decrease. Satisfying both of these first two points (high ON-state current and low OFF-state leakage) simultaneously implies that the access device needs to have a highly nonlinear I-V curve. However, it is difficult to extract this ratio for convenient comparison in a table because the voltages at which this ratio matters cannot be known until the NVM is specified.[25]

STT-MRAM devices require currents to flow in opposite direction during program and erase. Also, many RRAM device candidates show better cycling reliability and lower power requirements if operated in bipolar mode, where the program and erase operations are carried out under opposing voltage polarities. This implies that any access device that is used in conjunction with such memory elements must also be capable of **bidirectional operation**, as opposed to standard diodes. Given that program and erase currents typically have similar magnitudes, the access device should possess highly nonlinear I-V curves and drive high ON-state current density under both voltage polarities.

If, like most emerging resistive memory elements, the access device can be located above the wafer surface without consuming active Si wafer area, this frees up Si area for placement of some peripheral circuits underneath the crosspoint memory array. Such **process compatibility with 3D multilayer stacking** leads to more efficient utilization of the overall chip area for maximizing memory density. Furthermore, if the access device can be fabricated in a back-end-of-the-line (BEOL)-compatible process, this enables layering of multiple crosspoint arrays on top of each other.

Alternatively, vias and/or trenches deeply etched into a layered substrate can be recessed selectively and filled with memory material, creating a "Vertical-RRAM" structure. This approach eliminates costly lithography steps but

introduces the need for deep etches with extremely steep sidewalls,[31,32] much like the numerous bit-cost-scalable techniques being pursued for 3D NAND Flash.[33–35] More memory cells packed in a given chip area footprint leads to a lower effective cost per bit. BEOL-compatibility means that the maximum thermal budget during the access device fabrication process should not exceed ~400 °C; yet, the access devices themselves should be able to withstand up to ~400 °C for 2 h, in order to account for the processing needs of the memory elements and wiring levels fabricated above them.

Various resistive memory elements have different program/erase voltages and currents. It is important that the access device achieves **voltage compatibility with the memory element** with which it is paired. Access device turn-on voltages should have (or be tunable to) the appropriate values to ensure that the half-selected and unselected cells still have low leakage currents, even under array biasing conditions that provide the proper voltage drop across, and current through, the selected memory element for read and write conditions. If the access device's turn-on voltage is much lower than that needed across the selected memory element, array biasing voltages will lead to large currents through the half-selected cells, degrading the effective selectivity (ON/OFF ratio) of the access device. Typically, the voltage margin of the access device (defined as the voltage spread over which the leakage current is less than 10 nA) must be 0.3–0.5 V higher than the switching voltage of the NVM for 1000×1000 subarrays and ON-currents in the order of ~10 s of microamperes. These considerations must also accommodate variations in switching voltages due to device-to-device distributions and variations in operating temperature.

Since access device properties should not limit the overall memory chip performance and reliability, therefore **all other properties of the access device should be better than that of the memory element**. These properties—such as switching speed, cycling endurance, array yield, and variability—should be as good as or better than the corresponding properties of the memory element.

The set of requirements listed above make it very challenging to find access devices that are fully suitable for high-density, 3D multilayer resistive crosspoint memory. Section IV will survey a number of device candidates that have been proposed as access devices and highlight some of their strengths and weaknesses.

IV. TYPES OF ACCESS DEVICES

A. Si-based

Several silicon-based access devices have been proposed and demonstrated previously, leveraging the many decades of research, development, and manufacturing experience with silicon devices. Three-terminal devices such as transistors meet many of the access device requirements listed previously, including the ability to drive large ON-state currents (while maintaining a large ON-to-OFF current ratio), operate in bidirectional mode (although with some complications, due to the asymmetry introduced by the placement of the NVM cell), and the easy tunability of threshold voltage (with appropriate doping in the bulk, for instance). Moreover, the ability to use the transistor's gate voltage to constrain the saturation drain current offers an excellent way to enforce local current compliance (which is important for some NVM applications). For these reasons, and since silicon-based transistors can be mass-produced with very high yield and reliability, transistors are widely used as access devices in single-layer memory arrays.

However, transistor-based access devices also have their drawbacks, which include their relatively large cell size, complicated process, and incompatibility with 3D multilayer BEOL stacking (due to the required high processing thermal budget). While the latter two issues will likely be very difficult to solve, the use of vertical transistors (wherein current flows perpendicular to the Si wafer surface) offers a path to reducing the cell size. Wang *et al*. proposed a compact ($4F^2$ cell-size) FET access device based on vertical Si nanopillars with gate-all-around (GAA)[36] and experimentally demonstrated successful operation of HfO_2-based RRAM stacked above these transistors (Fig. 2). Alternatively, compact vertical bipolar junction transistors (BJTs) have also been used as access devices for resistive memories. Wang *et al*. showed switching of HfO_2-based RRAM stacked above vertical BJTs built in a standard commercial foundry CMOS logic process.[37] They pointed to the higher current driving capability of such vertical BJTs compared to planar MOSFETs built in the same CMOS process. Vertical BJT access

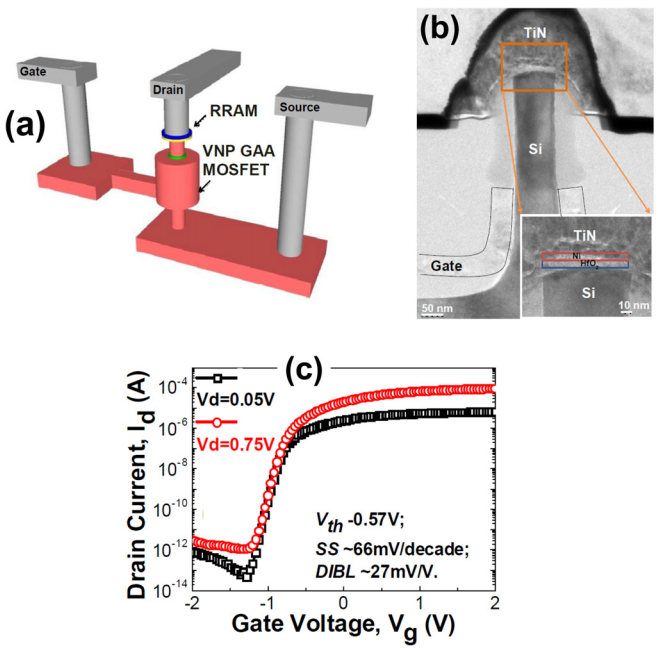


FIG. 2. (Color online) In a dense crosspoint 1T-1R array realized using vertical GAA Si nanopillar FETs, (a) each crosspoint intersection contains an RRAM element stacked above vertical Si transistor. (b) TEM cross section through experimental 1T-1R structure. (c) Drain-current as a function of gate voltage (for two different drain voltages) on a nanopillar GAA FET. Reprinted with permission from Wang *et al*., IEDM Tech. Dig. **2012**, 20.6. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.).

devices, capable of achieving very high drive currents (300 μA) and low base-emitter reverse bias leakage (<0.1 pA), have also been employed to drive PCM cells in a 45 nm generation 1 Gb chip.[38]

Two-terminal structures such as vertical Si PN junction diodes have been used as access devices for unipolar memory elements such as PCM. Oh *et al.* reported a 512 Mbit PCM chip[39] which had vertical Si PN junction diode access devices made using selective epitaxial growth (SEG) of Si over highly doped N-type Si substrate, using ion implantation to create superficial P-doped regions. They achieved a 5.8F^2 cell size at the 90 nm technology node. The reported I-V curves suggest very high ON-state current density (>25 MA/cm^2) and ON/OFF current ratio ~10^8. However, the area efficiency was somewhat degraded since contacts had to be made to the heavily *N*+ doped substrate after every eight cells in order to reduce the otherwise large series resistance that would be introduced by this layer. Also, second-order effects (due to vertical and lateral parasitic BJT currents) must be carefully considered in such structures, especially as the intercell spacing is decreased.[39,40] There have been numerous studies of SEG optimization[41,42] as well as alternative approaches based on solid-phase epitaxy[43] to build vertical Si PN diodes.

As an alternative to epitaxially grown Si diodes, polysilicon diodes have also been proposed. Sasago *et al.* showed polysilicon P-I-N diodes built above metal wiring[44] and reduced contact resistances to get high ON-state current density ~8 MA/cm^2. The worst-case forward/reverse rectification ratio was >~8×10^4. The authors were able to drive PCM cells with 160 μA RESET current through the polysilicon diodes. While the maximum process temperature was not specified in this paper, it is unlikely that these devices could be made in a typical sub-400 °C Cu-BEOL process while maintaining good enough crystalline quality and dopant activation for such high current density. In order to use polysilicon diodes for 3D-stacked multilayer crosspoint memory, one would need to either resort to high-temperature-compatible metallization schemes,[45] or employ processes such as laser annealing[46,47] for crystallization and activation of ion-implanted dopants in an effectively low thermal budget process.

Since some of the emerging memory technologies (such as RRAM or STT-MRAM) operate more efficiently in bipolar mode, two-terminal access devices capable of bidirectional operation have also been proposed. For instance, two-terminal NPN diodes have been proposed as access devices in simulation[48,49] and experimental studies.[50–52] The basic idea (Fig. 3) is to establish a NPN doping profile that results in the formation of a potential barrier for electron transport from one side of the device to the other at equilibrium (zero applied bias). Positive voltage applied to one of the electrodes will initially drop across the reverse-biased PN junction closer to that electrode. At higher voltages, the nonzero electric field within the middle P-region lowers the potential barrier for electron injection from the other (grounded) electrode. This mechanism is similar to the well-known "drain-induced barrier lowering" or "punch-through" effect that causes enhanced drain-to-source OFF-state leakage current in short channel MOSFETs. By symmetry, the same behavior would be seen if the opposite voltage polarity is applied across the NPN diode.

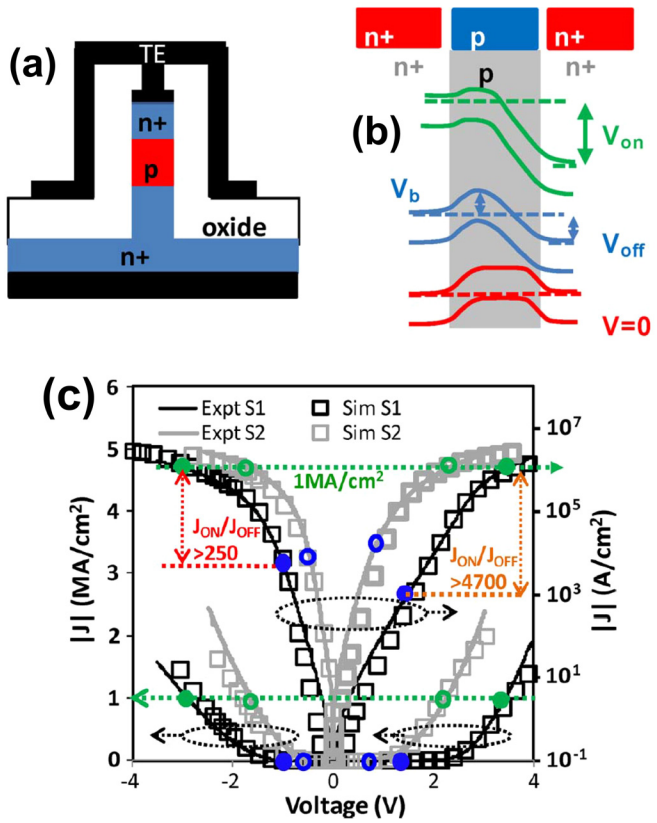


FIG. 3. (Color online) (a) Schematic cross-section through a two-terminal bidirectional NPN Si punchthrough diode selector. (b) Band diagrams through NPN structure at equilibrium (zero applied bias to positive electrode), during exponential turn-on (V_{off} applied bias), and in the ON-state (applied voltage V_{on}). Note that the barrier V_b for electron injection reduces with increasing applied voltage. (c) Experimental and simulated current density vs voltage plots for two structures, S1 and S2, with different doping profiles. Reprinted with permission from Srinivasan *et al.*, IEEE Electron Device Lett. **33**, 1396 (2012). Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

Srinivasan *et al.*[50] obtained >1 MA/cm^2 ON-state current density and maximum ON/OFF ratio > 10^6 (with OFF-current measured close to 0 V) on epitaxially grown Si NPN punchthrough diodes (Fig. 3). Threshold voltage tunability by adjusting the doping profile to obtain asymmetric turn-on voltages has also been shown.[51] Since the epitaxial growth process requires high temperatures (~700 °C, according to Ref. 50) as well as a crystalline Si template, such a device is unlikely to be either BEOL-compatible or easily stackable above the Si substrate. Polysilicon NPN structures made by Lee *et al.*[52] eliminate the need for a single-crystal Si seed layer but still require 800 °C annealing, again ruling out standard BEOL stacking. These devices showed bidirectional exponential turn-on with high maximum ON/OFF ratio but lower (~0.1 MA/cm^2) ON-state current density.

Park *et al.* studied a metal-semiconductor-metal (MSM) diode[53] with a P-type polysilicon layer contacted on both sides by metallic (Ta) electrodes which had assumably low

barriers to the Si conduction band edge. Simulations on such structures, proposed as access devices for STT-MRAM, showed >1 MA/cm^2 ON-state current density and >10^6 maximum ON/OFF ratio. The authors suggest that this structure offers potentially better control of the middle P-region thickness compared to the doped NPN devices. However, the I-V characteristics are dependent on the metal-semiconductor barrier height.

B. Oxide diodes

1. Oxide PN junction diodes

Most of the Si-based access devices discussed in Sec. IV A are not compatible with 3D-stacked multilayer crosspoint schemes mainly due to the high processing temperatures needed during their fabrication. This has motivated research into non-Si-based access devices which can be made at lower BEOL-compatible temperatures. Several groups have studied diodes made from semiconducting oxide heterojunctions.[54–58] Baek *et al*. fabricated PN diodes using P-type NiO and N-type TiO_2 at sub-300 °C processing temperatures[54] and demonstrated their operation with unipolar oxide RRAM elements. However, the ON-state current density of ~1 kA/cm^2 in these diodes was quite low.

Lee *et al*. built 2 stacked layers of small crosspoint arrays with unipolar NiO-based RRAM accessed through P-CuO_x/N-$InZnO_x$ diodes.[56] Their oxide diodes, deposited at room temperature, had higher ON-state current density (>10 kA/cm^2) than the P-NiO/N-TiO_2 diodes from Ref. 54, attributed to the lower bandgaps of these oxides. Kang *et al*. published more detailed studies of P-CuO_x/N-$InZnO_x$ diodes[57] including optical absorption measurements to estimate the bandgaps of the CuO_x and InZnO. Hall measurements revealed the conductivity of CuO_x to be P-type. Ahn *et al*. showed successful operation of one-time-programmable Al_2O_3 memory elements accessed through P-CuO/N-$InZnO_x$ thin film diodes and integrated in 8×8 crosspoint arrays.[58]

2. Metal-oxide Schottky barrier devices

Unipolar diodelike IV characteristics have also been reported in several two-terminal metal–insulator–metal (MIM) structures, where current rectification is due to transport across a Schottky-barrier at each metal-oxide interface.

Tallarida *et al*. obtained high rectification ratios (~10^7–10^8) and moderate ON-state current density (~10 kA/cm^2) in their devices with ZnO sandwiched between Ag and Ti/Au electrodes.[59] In these structures (of size down to 36 μm^2), ZnO was deposited by ALD at relatively low process temperatures (~100 °C) and intrinsically N-doped polycrystalline films were obtained. ZnO was assumed to form a Schottky junction with the Ag bottom electrode and an ohmic contact to the Ti/Au top electrode. The authors integrated these access devices with stacked unipolar NiO RRAM elements in 10 kbyte crosspoint arrays and reported RRAM SET switching through the diode.

Huang *et al*. measured rectification ratio >10^5 and ON-state current density ~2 kA/cm^2 in room-temperature-fabricated Ti/TiO_2/Pt MIM structures.[60] Based on the polarity of the I-V curves—low current and exponential turn-on with positive and negative voltage respectively applied to the Ti—the authors suggested, and confirmed via temperature-dependent measurements, that there were asymmetric barriers for electron transport from the metal electrodes into the TiO_2. The Schottky barrier height for electron injection from Pt into TiO_2 was found to be higher than that for injection from Ti into TiO_2. The role of inhomogeneous conduction through oxygen-deficient filaments in the TiO_2 was also highlighted. Reliability studies showed that the diodes survived ±3 V stress applied for 1000 s without degradation and 1000× repeated cycling between ±3 V with some reduction in forward and reverse currents. However, application of higher voltages (~5 V) with current compliance caused breakdown and turned these diodes into bipolar RRAM elements.

Park *et al*.[61] obtained very high rectification ratios ~10^9 in similar (Pt/TiO_2/Ti) MIM structures. Device operation was explained by assuming that the bottom Ti electrode interface was an ohmic contact while the top electrode Pt had a ~1 V barrier for electron transport into the TiO_2 deposited by atomic layer deposition (ALD). While the measured ON-state current density was very low (~10 A/cm^2) in these rather large devices (~60 000 μm^2), the authors used conducting AFM measurements to infer that the ohmic contact area at the Ti/TiO_2 interface was effectively confined to very small regions. Based on the estimated local ON-state current density of ~300 kA/cm^2 from these measurements, the authors suggested that these devices would show better ON-state current density when scaled down in size. Measurements on smaller crosspoint-type Pt/TiO_2/Ti/Pt devices[62] indeed showed higher ON-state current density ~300 kA/cm^2 in 2×2 devices, although with lower rectification ratio (~2.4×10^6) compared to the earlier devices in Ref. 61.

There have been reports of Ni/TiO_2/Ni structures,[63,64] which show bipolar nonlinear I-V characteristics. The devices published by Huang *et al*.[64] had ON/OFF ratio of ~10^6 between 0 and ±2 V. Furthermore, the nonlinearity was maintained even under high temperature (125 °C) operation and after 1000 × DC cycling. The device operation was ascribed to electron emission over the Ni/TiO_2 Schottky barrier whose height was estimated to be ~0.58 V at zero applied bias. Ni/TiO_2/Ni devices of varying sizes down to 0.36 μm^2 were reported by the same group in Ref. 63.

Figure 4 depicts some measured I-V curves from this work.[63] The authors suggested that nonuniform current conduction explains the observed weak scaling of ON-state current with device area. They found that the calculated ON-state current density increased after reducing the device area and obtained ~100 kA/cm^2 in their smallest (0.36 μm^2) devices. These Ni/TiO_2/Ni structures were used as bidirectional access devices to access Ni/HfO_2/Pt bipolar RRAM. Leveraging the room temperature deposition process for the access devices as well as RRAM elements, the authors built small 8×8 crosspoint 1S-1R arrays on flexible plastic substrates.

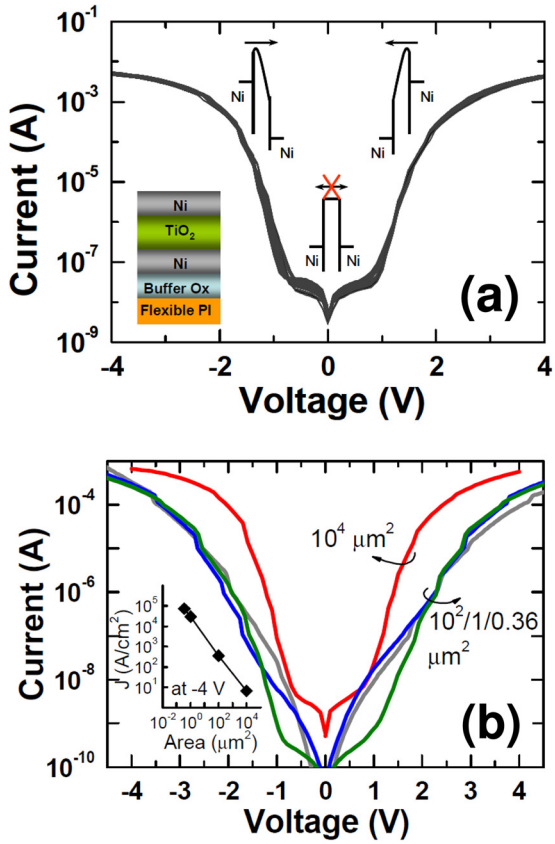


FIG. 4. (Color online) (a) Bidirectional exponential current-voltage curves measured on experimental Ni/TiO_2/Ni devices. (b) Experimental I-V curves measured on Ni/TiO_2/Ni bipolar selectors of varying area. The inset shows that the ON-state current density increases as the device area is reduced. Reprinted with permission from Huang *et al.*, IEDM Tech. Dig. **2011**, 31.7. Copyright 2011 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

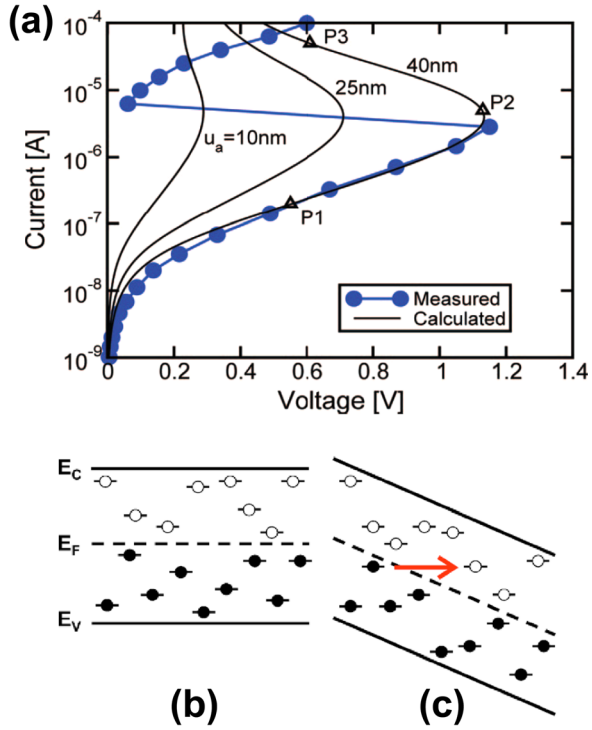


FIG. 5. (Color online) (a) Measured I-V curve for a phase-change memory cell with amorphous $Ge_2Sb_2Te_5$ chalcogenide and dependence of threshold voltage on thickness of the amorphous chalcogenide. Reprinted with permission from D. Ielmini, Phys. Rev. B **78**, 035308 (2008). Copyright 2008 The American Physical Society. (b) Electron tunneling, from lower energy filled traps into higher energy empty traps, can occur under high field (Ref. 67). Reprinted with permission from D. Ielmini and Y. G. Zhang, J. Appl. Phys. **102**, 054517 (2007). Copyright 2007 AIP Publishing LLC.

C. Threshold switch

A threshold switch makes for an ideal access device, since from an initial highly resistive state the device turns ON to a highly conducting state as soon as a threshold voltage or current is applied. Various types of threshold switches such as ovonic threshold switches, metal–insulator-transition (MIT)-based devices, and threshold vacuum switches have been proposed for application as access devices for crosspoint memory.

1. Ovonic threshold switching

The ovonic threshold switching (OTS) phenomenon was first reported by Stanford Ovshinsky in thin films of amorphous chalcogenide alloys.[65] Unlike high field electrical breakdown phenomena, ovonic threshold switching is repeatable and nondestructive. The current–voltage (I-V) curve of a simple two terminal threshold switching device measured with current sweep mode is generally "S" shaped [Fig. 5(a)], switching rapidly from an initial high resistance "OFF" state into a highly conductive dynamic "ON" state once the applied voltage exceeds the *threshold voltage*, V_{th}. The device remains in the ON state so long as a minimum holding voltage, V_h, (or current I_h) is present across (through) the device, otherwise switching back into the OFF state.

This threshold switching occurs after a finite delay time (t_d),[68] which could potentially affect the upper limit for the speed of the integrated memory device.[69] However, it has been found that the switching delay time decreases with increasing applied voltage and even ~ns switching speed can be obtained.[68–71] The threshold voltage tends to decrease with decreasing thickness,[72,73] as well as with increasing temperature,[73] suggesting that the threshold event is triggered by a threshold electric field.[72,73] On the other hand, the holding voltage is independent of the chalcogenide thickness but is affected by the electrode materials.[73]

To explain the threshold switching phenomenon, various theoretical models including thermally induced instabilities,[74] impact ionization and Shockley–Hall–Read recombination,[68,75] polaron instabilities,[76] and field induced crystallization[70,77] have been proposed.

Recently, Ielmini *et al.* proposed a model based on trap-limited current conduction, similar to Poole–Frenkel transport at high electric fields.[66,67] Applied electric field decreases the barrier height for electron hopping conduction, which in turn dramatically increases conductivity.[66,67] As the applied voltage increases, the carriers gain energy and tunnel into higher energy traps [Fig. 5(b)]. When these traps are filled, mobility increases and current rises until switching occurs. This model has successfully explained

experimentally observed trends for threshold voltage as a function of both temperature and thickness.[66]

For use as an access device, it is important that any undesired structural rearrangement, such as crystallization, is suppressed during threshold switching (and high-current operation) of an OTS device.[73] While similar chalcogenide materials are used in phase change memory (PCM) devices, where both threshold switching and fast crystallization is desirable, it is possible to find materials that exhibit threshold switching without fast crystallization.[68,69,71,73] Various materials have been reported as ovonic threshold switching–based memory access devices.[69,78,79]

Kau *et al*. demonstrated the successful integration of an OTS access device on top of a "mushroom-cell" PCM device in a stackable crosspoint structure.[78] Figure 6(a) shows a TEM of the PCM device and OTS access device, while Fig. 6(b) shows the measured I-V characteristics. Unfortunately, the use of a linear current scale makes it impossible to gauge the actual ON–OFF contrast of this device.

The higher threshold voltage of the OFF state of the stacked device pair is attributed to the serial connection of two amorphous alloys, OTS and amorphous PCM.[78] Figure 6(c) shows the cut-away SEM image of a 64 Mb "PCMS" (PCM + OTS) test chip fabricated using 90 nm technology node,[78] which was used to explore cell-sizes ranging from 40 to 230 nm. Single PCMS devices showed fast reset speed of 9 ns and pulse endurance of 10^6 cycles; statistics over a 2 Mb block showed more than 1 V of dynamic range between the highest threshold voltage of a PCMS device in the SET state and the lowest threshold voltage of a device in the RESET state.[78]

Lee *et al*. have reported threshold switching in AsTeGeSi based material by utilizing electronic charge injection.[79] Improved characteristics, including a reduction in threshold voltage and current, were observed after N_2 plasma nitridization, which led to higher tellurium (Te) composition. Fabricated AsTeGeSiN threshold switches have shown pulse switching endurance of more than 10^8 cycles at a pulse width of 1 μs. Such an OTS device (1S) was integrated with TaO_x based resistive memory (1R) to realize a 1S1R memory cell. The 30 nm thick (10 nm of Ta_2O_5 on 20 nm TaO_x) RRAM cell was located under the 40 nm thick OTS cell, separated by the bottom TiN electrode of the OTS deposited onto the top Pt electrode of the RRAM cell. An ON-OFF selectivity of 100 was observed for 30 μm devices with a leakage floor of $\sim$1 μA for $V_{SET}/2$, but scaling experiments suggested that a selectivity of 1000 could potentially be achieved for 30 nm devices.

2. Metal–insulator transition

MIT behavior in transition metal oxides has been a widely researched topic in the field of condensed matter

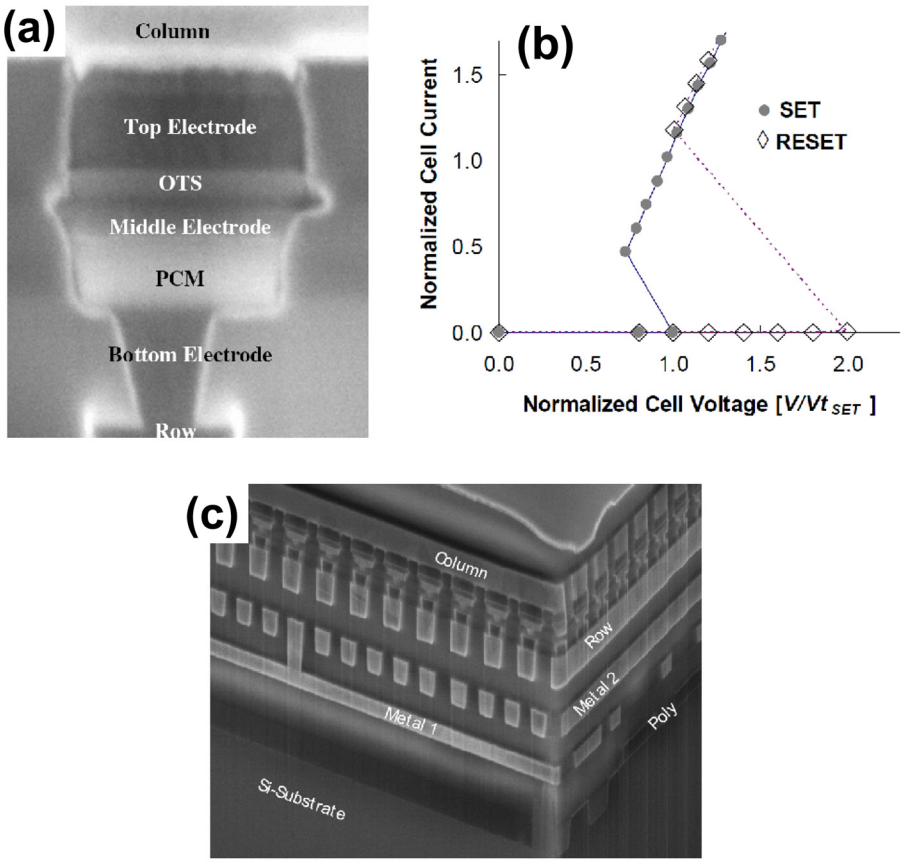


FIG. 6. (Color online) (a) SEM cross section of a PCMS (OTS + PCM) cell; (b) I-V Characteristics of a such a cell in SET (On state) and RESET (Off state) state; and (c) one layer of an integrated PCMS array. Reprinted with permission from D. C. Kau *et al*., IEDM Tech. Dig. **2009**, 27.1. Copyright 2009 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

materials.[80] It has been observed that various stimuli (magnetic, optical, and pressure) can activate MIT behavior.[80] Recently, electrically or thermoelectrically triggered MIT devices have become attractive candidates as bidirectional access devices for ReRAM application because of their simple metal/metal–insulator stack, high ON/OFF ratio, and fast transition speed.

Vanadium dioxide VO_2 possesses a first-order metal–insulator transition. Although its XRD pattern exhibits only a single monoclinic phase, MIT behavior cannot be observed in microscale VO_2 device—in such large devices, there are invariably numerous leakage current paths through nonstoichiometric oxide defects.[81,82] In nanoscale devices, however, as applied voltage is increased from 0 to 0.5 V, the current suddenly increases at 0.35 V as the electrical property of VO_2 changes from an insulating to a metallic state [Fig. 7(a)].[81,82] When the voltage is brought back down below 0.2 V, the current suddenly decreases as the VO_2 returns to its initial insulating state. Nanoscale VO_2 devices show very fast response (<20 ns) to an input voltage signal, and individual devices show stable and uniform threshold and hold voltages over 100 successive switching cycles [Fig. 7(b)].[81,82]

While VO_2 exhibits excellent MIT characteristics at room temperature, its transition temperature is too low (~67 °C) for practical device application at operating temperatures up to 85 °C. NbO_2, however, offers a bulk transition temperature of 1081 K.[83] Recently, ultrathin NbO_2 film, intentionally fabricated by reactive sputtering method and patterned to nanoscale device area, has been investigated for access device applications.[84,85] Similar to VO_2 devices, individual NbO_2 MIT-based access devices exhibit well-behaved and repeatable hysteretic I-V characteristics over 1000 switching cycles (Fig. 8).[84] Thermal simulations have shown that Joule heating through the device causes the internal maximum temperature to exceed 1100 K, triggering the metal–insulator transition.[84] Unlike devices using VO_2, NbO_2-based access devices show excellent thermal stability up to 160 °C, sufficient for reliable memory application.[84] However, the threshold voltages are significantly larger and the hysteretic window smaller than VO_2, and the leakage current corresponding to the half-select condition is still fairly high.

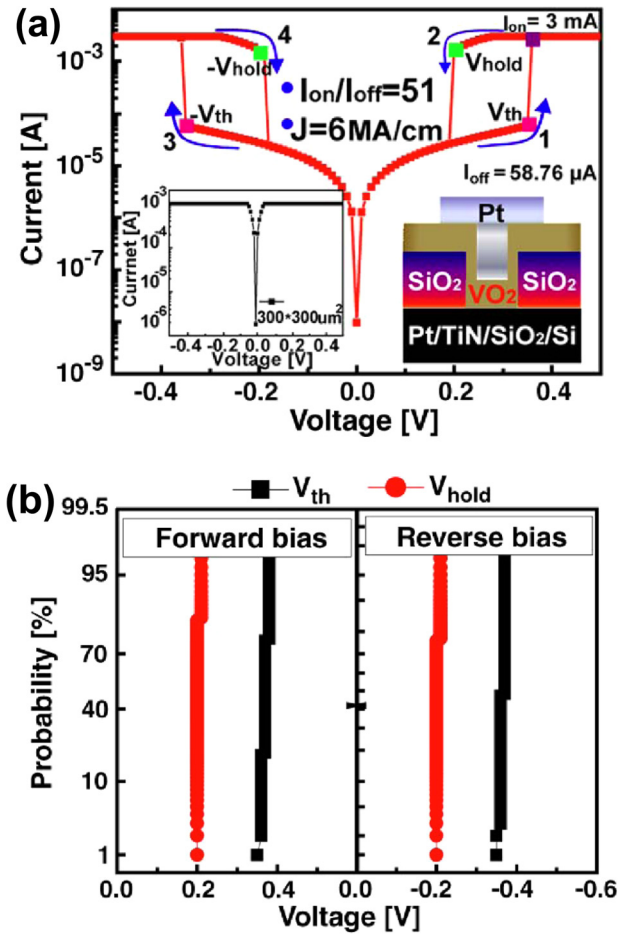


FIG. 7. (Color online) (a) I-V characteristics of a nanoscale vanadium dioxide device. (Inset) Left: I-V characteristics of microscale device, right: structure of nanoscale Pt/VO_2/Pt device. (b) Distribution of V_{th} and V_{hold} over 100 successive switching cycles. Reprinted with permission from Son *et al.*, IEEE Electron Device Lett. **32**, 1579 (2011). Copyright 2011 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

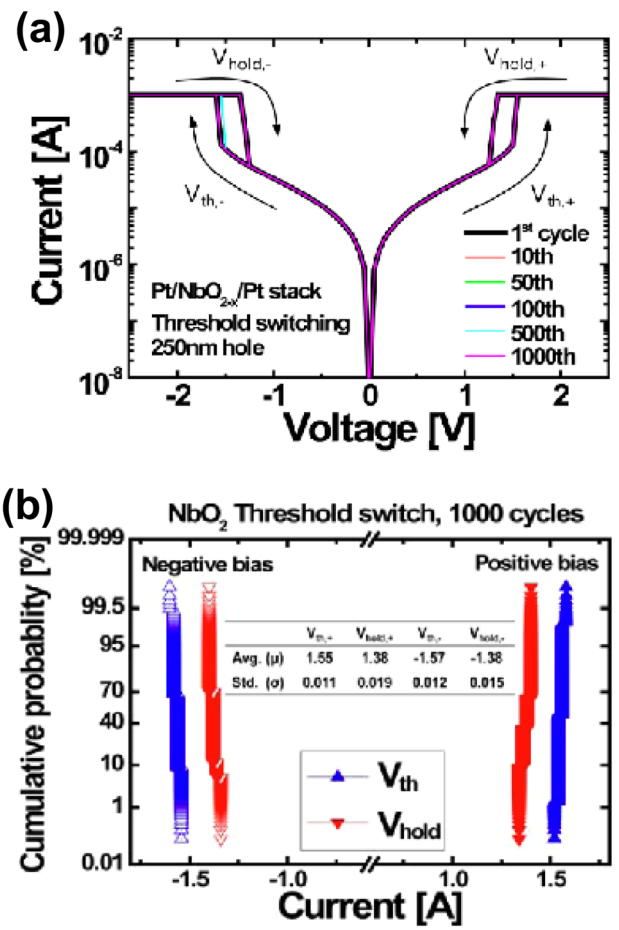


FIG. 8. (Color online) (a) I-V characteristics of MIT behavior of Pt/NbO_2/Pt device. (b) Distribution of V_{th} and V_{hold} during single-device cycling. Reprinted with permission from S. Kim *et al.*, Symp. VLSI Technol. **2012**, T18.3. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

3. Threshold vacuum switch

Recently, a threshold vacuum switch (TVS) was proposed as a bidirectional access device for 3D stackable crosspoint arrays.[86] The TVS stack consists of W/TiN bottom electrode, TiN/vacuum top electrode with vacuum gap of ~1 nm (established by a SiN sacrificial layer), and WO_x as a switching layer, prepared by electrochemical oxidation.

Volatile switching behavior with high current density (>100 MA/cm^2) and high ON-OFF selectivity of $>10^5$ between ON and OFF states was observed in individual devices upon application of modest (1.5 V) dc or pulsed bias. The TVS ON state in the low voltage region corresponds to direct tunneling, with the high voltage region limited by the ohmic conduction WO_x layer. The TVS OFF state is described as the superposition of direct tunneling and nonlinear conduction. Constant voltage stress tests with ±2 V applied, resulting in high (100 MA/cm^2) current density, were made, suggesting an equivalent endurance of over of $>10^8$ cycles using 100 ns programming pulses. However, it is not clear how easily an array of devices with identical ultranarrow vacuum gaps could be fabricated.

D. Oxide tunnel barrier

One attractive approach for achieving strongly nonlinear I-V characteristics is to use a thin oxide or nitride layer as a tunneling barrier. Such field-induced tunneling behavior is the basis of programming in conventional NAND FLASH memory,[87] and allows currents to increase exponentially by several orders of magnitude at both polarities via electron tunneling behavior. Tunnel barriers can be formed by various high-k materials (HfO_2, SiO_2, ZrO_2, and TiO_2)[88,89] and the field sensitivity enhanced by tunnel-barrier engineering of multilayer oxides. According to how the dielectrics are stacked, such engineered tunnel barriers are classified as crested and variable oxide thickness (VARIOT)-type barriers.[90,91]

1. Single layer oxide-(nitride-)based access devices (TiO_2 and SiN_x)

By choosing electrode materials with suitable workfunctions relative to the oxide/nitride-based semiconductor layers (TiO_x or SiN_x), Schottky barriers can be formed at both interfaces (Fig. 9).[92] At low bias, electron injection is initially limited by one of the two barriers. When higher bias is applied, band-bending leads to increased current.

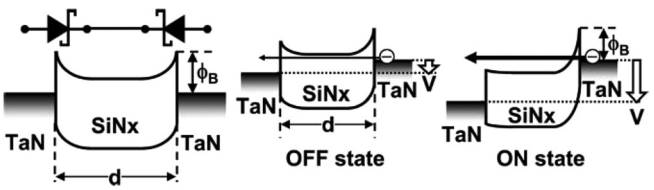


FIG. 9. Schematic diagram of a tunnel barrier device composed of two back-to-back Schottky barriers. Reprinted with permission from Kawahara *et al.*, IEEE J. Solid-State Circuits **48**, 178 (2013). Copyright 2013 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

Symmetrical device structures lead to the symmetric I-V curves needed for bipolar RRAM applications.

In devices composed of an ultrathin TiO_2 layer deposited by atomic layer deposition (ALD) on a TiN bottom layer, when a positive voltage is applied to the Pt top electrode, electrons transfer from TiN layer to Pt layer.[93] In this case, current density is relatively high (15 kA/cm^2 at 2 V) because of the small (~0.4 eV) Schottky barrier at the TiN/TiO_2 interface.[93] The current density under negative bias is smaller (1 kA/cm^2 at −2 V) because the barrier at the Pt/TiO_2 interface is larger. The Schottky barrier height and the film thickness are both parameters that can increase the current density exponentially. By changing the injection time of oxidizer in the TiO_2-ALD process, the ratio of lattice to non-lattice oxygen can be controlled, leading to a change in the number of oxygen vacancies and a modulation in the work-function of TiO_2.[93] This allows the current density level of access devices to be adjusted to match with the particular ReRAM device of interest.

2. Multilayer oxide-based access devices

Another method for achieving a strongly nonlinear I-V curve is tunnel-barrier engineering of a multilayer oxide stack, in which the stacking of the tunnel dielectrics is optimized to enhance the field sensitivity and thus strengthen the nonlinearity in the I-V curve. The engineered tunnel barriers can be classified as crested,[89,90,94] in which the tunneling barrier height is maximum in the middle and gradually decreases toward the conducting electrode; and VARIOT-type, where the central region has a lower bandgap than the outer regions.[95] VARIOT-type stacking has been realized by using simple thermal oxidation process on TaO_x/TiO_2 stack which resulted the formation of a Ta_2O_5 tunneling barrier as shown in Figs. 10(b) and 10(c).[91,96]

High current densities (>10 MA/cm^2) and a nearly symmetric I-V curve have been observed, together with a large (~10^4) ON-OFF contrast [Fig. 10(a)]. In the low bias OFF state, electron conduction is restricted by the higher total thickness. When the bands bend at higher bias, current increases exponentially under either polarity, by tunneling through the extremely thin oxide layer. No disturbance was observed under V_{READ} and $V_{READ}/2$ conditions out to 10^{10} pulses of 100 ns width.[91]

A scalable Cu-based conductive bridging RAM (CBRAM) was directly integrated on top of such an access device, using a via-hole structure [Fig. 11(a)].[91] Figure 11(b) shows the I-V characteristics of the 1S-1R combined device pair. The presence of the access device suppresses the current at the original $V_{READ}/2$ by a factor of 10^3.

E. Mixed-ionic-electronic-conduction

Copper-containing MIEC materials[28,97–100] have become an intriguing choice as 3D-ready access devices for NVM. MIEC materials can be processed at sub-400 °C temperatures making them back-end-of-line (BEOL) friendly. MIEC-based access devices offer large ON/OFF ratios ($>10^7$), high voltage margin V_m (over which current <10 nA), and

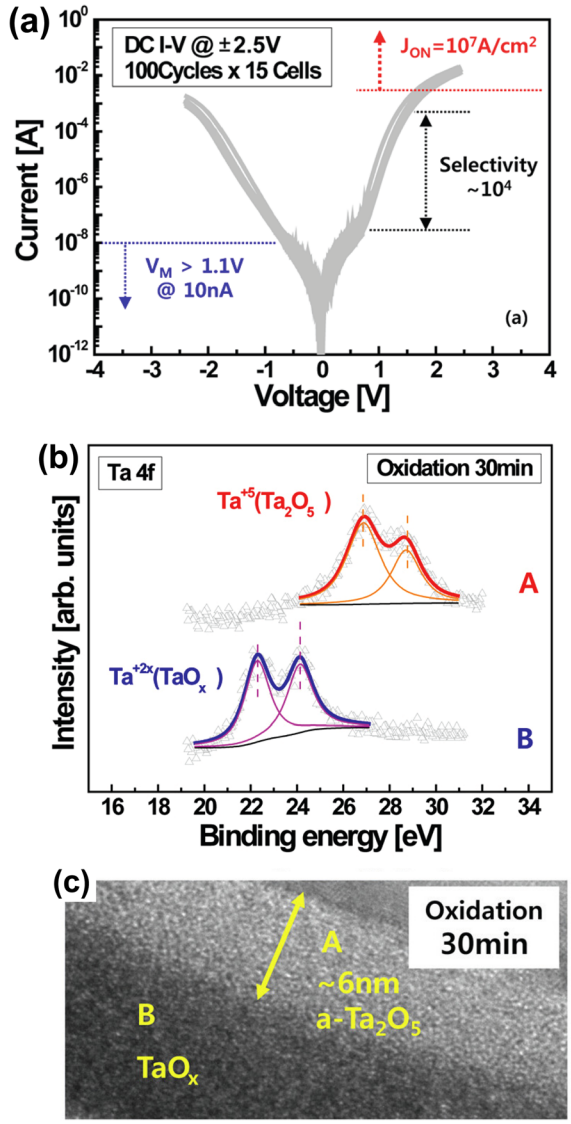


FIG. 10. (Color online) (a) Typical I-V switching characteristics of 15 cells with 100 cycles, using engineered tunnel barriers ($Pt/Ta_2O_5/TaO_x/TiO_2/Pt$ stack); (b) X-ray photoelectron spectroscopy (XPS) spectra and (c) TEM images after 30 min oxidation. In the top electrode interface region (A), the main composition is Ta_2O_5, an insulator, and in the bulk region (B), suboxides such as TaO_x. Reprinted with permission from Woo *et al.*, Symp. VLSI Technol. **2013**, T12.4. Copyright 2013. The Japan Society of Applied Physics.

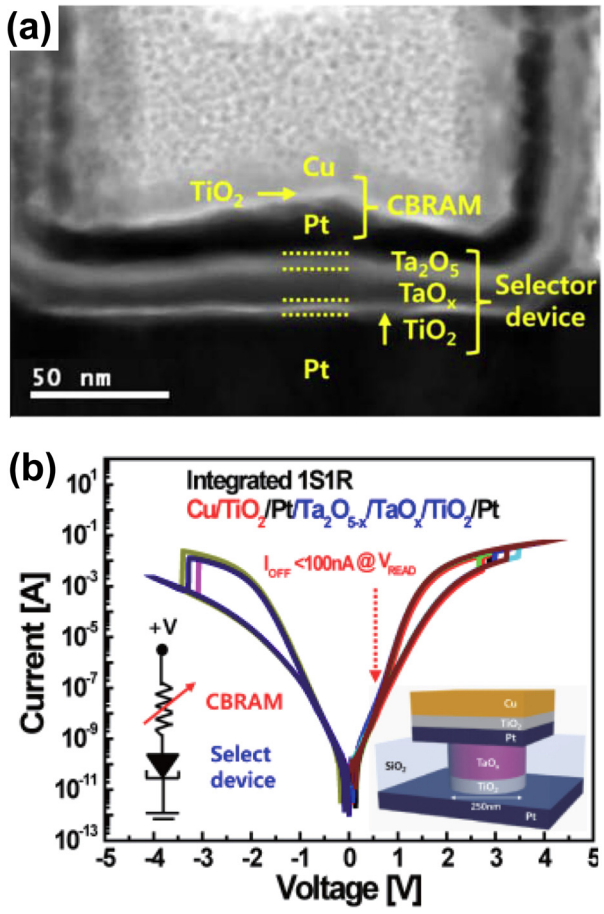


FIG. 11. (Color online) (a) Cross-sectional TEM image of the selector integrated with CB-RAM in the 250-nm via-hole structure and (b) the I-V characteristics of CB-RAM and integrated 1S-1R device. Reprinted with permission from Woo *et al.*, Symp. VLSI Technol. **2013**, T12.4. Copyright 2013 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.).

ultralow leakage (<10 pA). In pulse-mode, the devices can carry high current densities needed for PCM and fully bipolar operation needed for high-performance RRAM.[97,98] MIEC materials contain a large amount of mobile copper ions, which can move readily within devices made from MIEC-materials. At low bias, Schottky barriers at the MIEC–electrode interfaces are believed to strongly suppress current flow, leading to the low leakage. As bias increases in either direction, copper ions move toward the negatively biased electrode and vacancies toward the positively biased electrode within the device, modulating the interfaces and leading to an exponential increase in electronic current. Although the current eventually saturates, extremely high current densities have been observed [up to 50 MA/cm^2 (Ref. 97)]. Although higher drive currents have been observed to sharply reduce cycling endurance, ~10^8 cycles have been demonstrated at ~150 μA.[98] No sudden breakdown has been observed, but practical maximum operating voltages are constrained by rapid endurance failure as currents increase above ~150 μA.[98]

An integration scheme has been developed for large scale manufacture of MIEC-based access devices and demonstrated on 8 in. wafers. Major steps in the scheme involve deposition of Cu-containing MIEC materials onto patterned vias followed by an optimized CMP process;[98] the vias are finally capped with a patterned nonionizable top electrode. The devices can then be tested either using a probe-tester or using conductive atomic force microscopy (C-AFM). With C-AFM methodology, MIEC devices fabricated in a short-loop with minimal wiring, yet the three critical elements of the basic device structure can be evaluated: the bottom electrode, the MIEC-based material deposited and polished in a confined via, and a patterned top electrode. C-AFM testing has validated single-target deposition and revealed a wide processing temperature window (up to 500 °C) for these MIEC-based access devices.[28] Although these access

devices themselves do not need high processing temperatures, MIEC-based access devices would not be adversely affected if a cointegrated RRAM or other NVM were to require such processing conditions.

In order to test large numbers of MIEC-based access devices, bidirectional array diagnostic monitor arrays of up to 512×1024 integrated MIEC access devices were tested using integrated 1-bit sense-amplifiers and a fast electrical tester (Magnum 2EV).[28] Cumulative distribution functions (CDFs) of the bitline voltage V_{BL} needed to produce various device currents I_d show tightly distributed array I-V characteristics [Fig. 12(a)]. All 524 288 MIEC devices—100%—had $V_m > 1.1$ V, and 99.955% of access devices fell within ± 150 mV of the median voltage margin $V_m = 1.36$ V at 10 nA [Fig. 12(c)].[28]

C-AFM studies were used to explore scaling of MIEC-based access devices in both vertical and lateral dimensions. As MIEC-based access devices become thinner (at nominally similar lateral dimensions), the voltage margin remained mostly unchanged until the minimum gap between electrodes, $d_{\min}$, reaches ~11–12 nm.[99] When the lateral dimension is also made small, leading to highly scaled vias, these MIEC-based devices show both high yield and high voltage margin.[99] In order to enable pulse testing of ultrascaled vias, a bottom contact was fabricated a few micrometers away from the devices. Despite their small size, these MIEC access devices can still rapidly drive the large currents needed for NVM switching [Fig. 13(a)]. Voltage margin V_m (at 10 nA) improves markedly as devices are scaled in lateral size. Aggressively scaled MIEC access devices retain all requisite characteristics, including low leakage (<10 pA) and the large voltage margins ($V_m > 1.50$ V) needed for large arrays, even

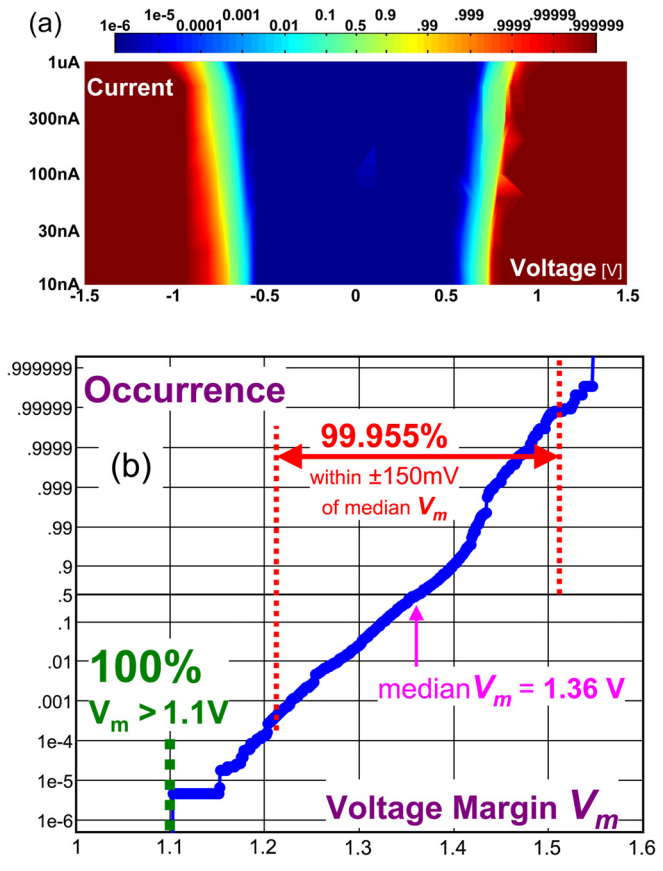


FIG. 12. (Color online) CDFs across bitline voltage V_{BL} at various device currents I_d (a) show array level I-V results with tight distributions across a 512×1024 array of integrated MIEC access devices Ref. 28. (b) Within this same 512×1024 array, there were no leaky devices; 100% of the array showed $V_m > 1.1$ V, while 99.955% of MIEC access devices had voltage margins V_m at 10 nA within ± 150 mV of the median of 1.36 V. Reprinted with permission from Burr *et al.*, Symp. VLSI Technol. **2012**, T5.4. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

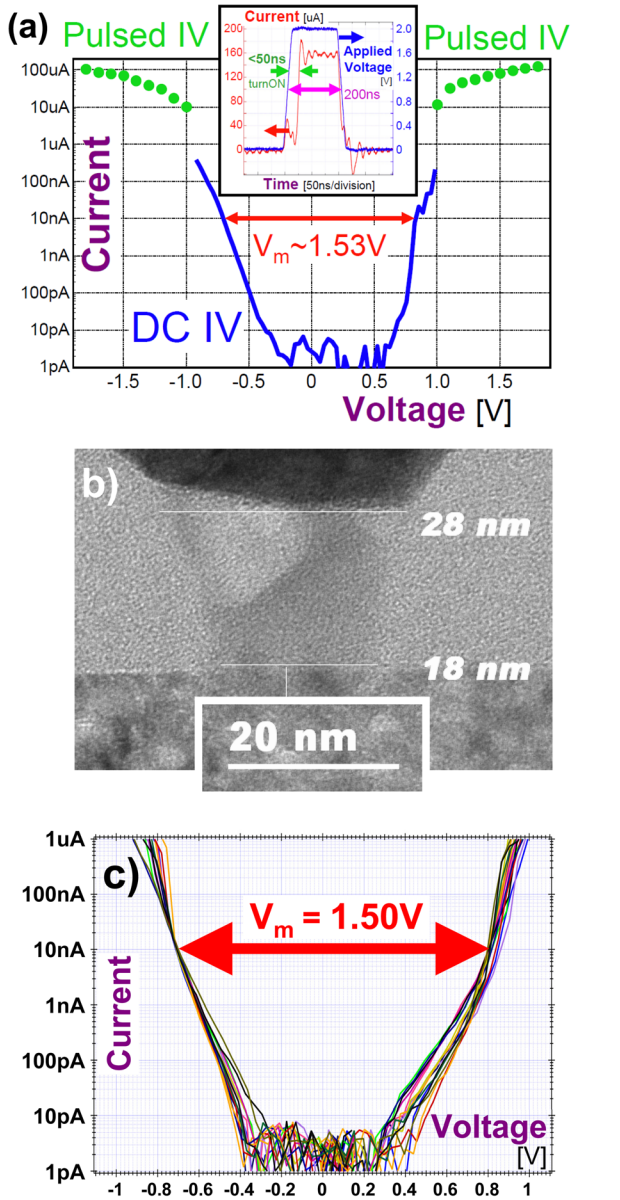


FIG. 13. (Color online) (a) In addition to high yield, scaled short-loop MIEC access devices (Ref. 99) exhibit the same $>10^7$ ON-OFF contrast, <50 ns turn-ON times (test-setup-limited), and low leakage shown in larger devices (Refs. 28 and 98). Aggressively scaled short-loop MIEC access devices, with (b) both TEC and BEC <30 nm, also show (c) large voltage margins and low leakage. Reprinted with permission from Virwani *et al.*, IEDM Tech. Dig. **2012**, 2.7. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

with both top and bottom critical dimensions (CDs) <30 nm [Figs. 13(b) and 13(c)]. Unlike thickness scaling, where leakage increases sharply below $d_{min} \sim 11$nm, no lower limit to CD scaling has yet been identified.[99]

To demonstrate access characteristics of MIEC devices, an integration scheme whereby MIEC access devices were fabricated immediately above small (CD ~ 35 nm) PCM devices to form PCM + MIEC device pairs was successfully achieved. Figure 14 demonstrates that PCM + MIEC device pairs show endurance in excess of 100 000 cycles, despite the repeated application of RESET pulses >200 μA and 5 μs long SET pulses at ~90 μA.[28] This pairing also allowed demonstration of NVM write speed capabilities using an MIEC-based access device to rapidly RESET a cointegrated PCM device. After each 15 ns RESET pulse at varying amplitude, a long SET pulse recrystallized the doped-$Ge_2Sb_2Te_5$ PCM material. After each RESET and SET pulse, bipolar DC IV curves [Fig. 15(a)] showed the expected change in the PCM resistance above the same low-leakage characteristics of the MIEC access device.[28] The application of shaped pulses or a transient "overvoltage" read readily allows MIEC access devices to support ~5–10 μA NVM reads in <50 ns.[99,100]

In the context of a large crosspoint array, MIEC-based devices must rapidly change their role depending on whether their cell is in the *unselected*, *selected-for-read*, *selected-for-write*, or *half-selected* state. It has been shown that the low leakage offered by MIEC-based access devices representing the *unselected* state can be held for hours [Fig. 15(b)], while higher *half-selected* leakage can be sustained for at least a few seconds, more than long enough for millions of successive 1 μs reads to the same bitline.[100]

Despite relying upon Cu-ion motion to achieve their highly nonlinear characteristics, MIEC-based access devices can return rapidly to low leakage after leaving either the *selected-for-read* or *selected-for-write* states.[100] The leakage recovery after write (30–50 μA) operations requires ~1 μs spent at zero voltage bias, while postread (3–6 μA) recovery is even faster. Recovery can be accelerated by application of opposite polarity rather than just zero voltage bias. Read operations can be sub-50 ns, fast enough for use with MRAM [Fig. 15(c)][100] using transient "overvoltage" schemes. However, application of large overvoltages can be problematic within large arrays because of the presence of large capacitances. Fortunately, thickness scaling allows thin MIEC access devices to offer similar speeds while requiring only modest overvoltages, reducing the possibility of read disturb.[100]

MIEC-based access devices are well-suited for both the scaled CDs and thicknesses of advanced technology nodes and for the fast read and write speeds of emerging NVM devices. Future improvements would include a better quantitative understanding of the interaction between the mobile Cu-ion dopants and the resulting electronic current,[101] improvements in endurance at high current [currently ~10^8 cycles at 150 μA (Ref. 98)], verification of high endurance in scaled devices at both high and low operating temperatures, and increases in the voltage margin so as to enable large arrays for NVM devices requiring >1.5V switching voltages.[25]

V. SELF-SELECTED RESISTIVE MEMORY

Any memory device whose IV characteristics are highly nonlinear in all resistance states can be implicitly "self-selective," thus obviating the need for an explicit access device. Such inherent device IV characteristics may sufficiently reduce sneak path leakage currents, leading to acceptable read margin and sufficiently low excess power dissipation in nonselected devices during write.

A number of attractive device configurations have been integrated to realize self-selected resistive memory. As an example, in a complementary resistive switch (CRS), two bipolar RRAM cells (preferably with symmetric set/reset

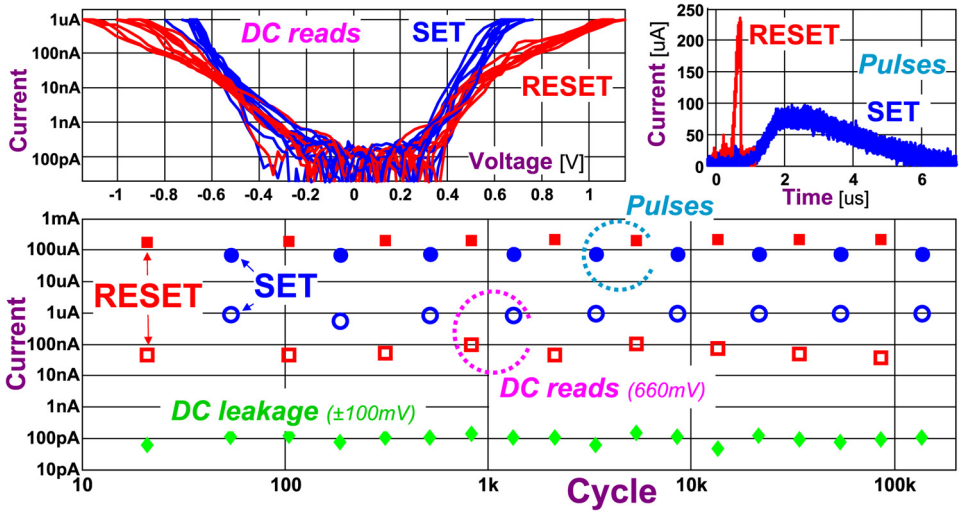


FIG. 14. (Color online) Endurance of an integrated PCM + MIEC device-pair to >100 k cycles, with RESET currents >200 μA and 5 μs-long SET pulses (~90 μA). No access device degradation or PCM failure had occurred at the time testing was terminated. Reprinted with permission from Burr *et al.*, Symp. VLSI Technol. **2012**, T5.4. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

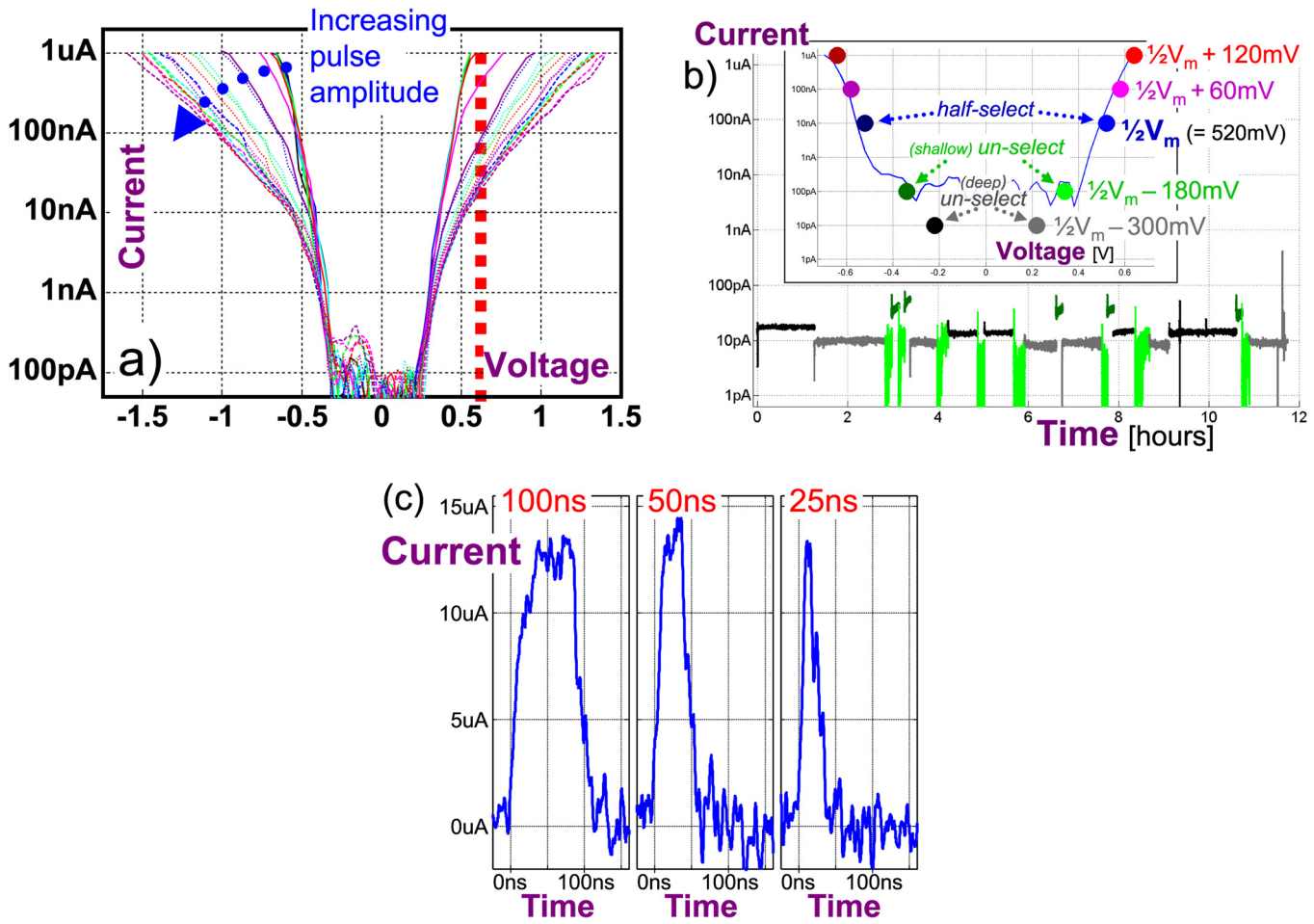


FIG. 15. (Color online) (a) After each single 15 ns RESET pulse, bipolar dc IV curves show the large resistance contrast (~1 MΩ) of PCM RESET devices, yet the low-leakage characteristics of the MIEC access device remain unaffected (Ref. 99). MIEC access devices (b) maintain low-leakage over hours of exposure, whether in a deep (±230 mV) or shallow *unselect* (±350 mV) condition, and (c) can transition from *half-select* to 10 μA read currents in ~50 ns. Reprinted with permission from Burr *et al.*, Symp. VLSI Technol. **2013**, T6.4. Copyright 2013 The Japan Society of Applied Physics.

voltages) are stacked back to back with one common electrode. Another example is a hybrid RRAM cell, wherein built-in access device functionality is integrated in a single cell structure using one of the standalone access device mechanisms already discussed. In a nonlinear RRAM cell, additional nonlinearity is introduced into the LRS by inserting barrier layers at the electrode/switching material interface to reduce the leakage current. The desired outcome in each of these approaches is access device functionality with a simplified fabrication process and smaller cell size and/or thickness.

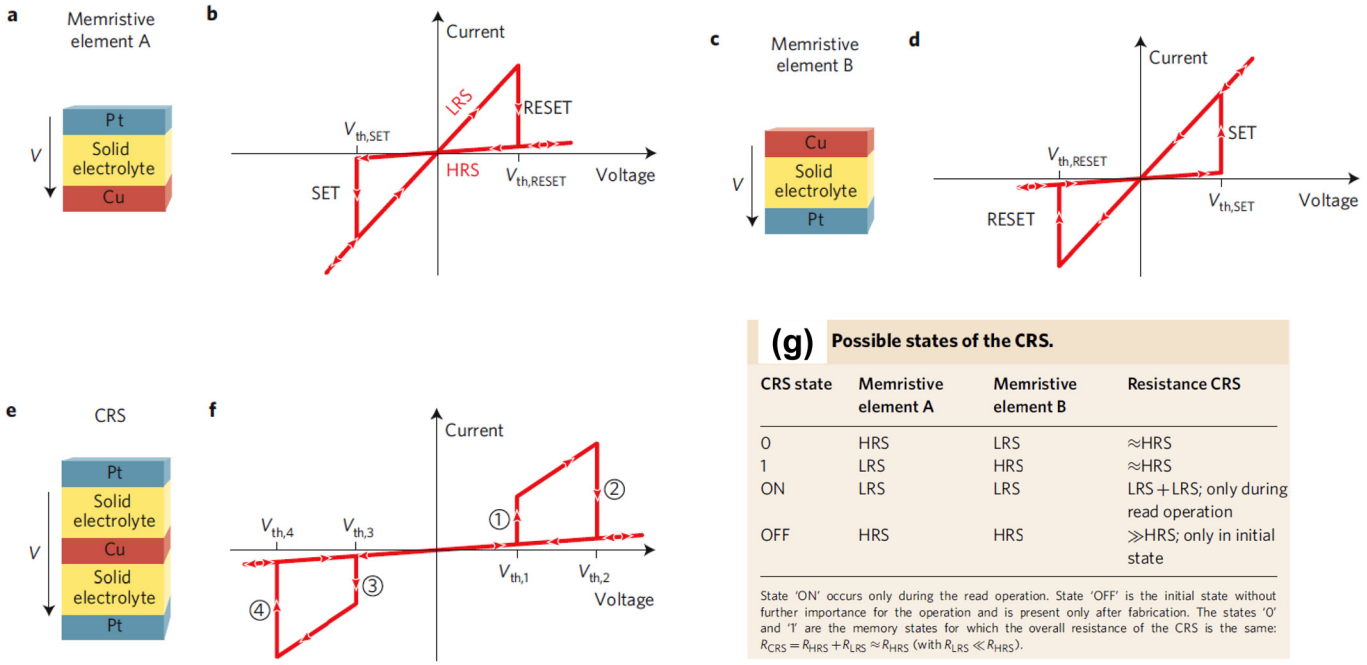
A. Complementary resistive switch

The CRS proposed by Linn *et al.*[102] consists of two antiserially connected bipolar resistive switches, as illustrated in Fig. 16. In this first manifestation, each resistive switching element was a conductive bridging memory (CBRAM) with an oxidizable (Cu) electrode and a solid electrolyte (such as GeSe). A positive voltage on the TE forms a filament in the bottom element and dissolves any filament in the top element.

The four possible combinations of HRS (or no filament) and LRS (or copper filament) are summarized in Fig. 16(g).[102] After initialization of the original HRS/HRS state, HRS/LRS can be defined as the state "0," and LRS/HRS as the state "1." The fourth state, LRS/LRS, is the "ON" state of the CRS cell obtained during the reading operation. Sneak current is suppressed by the presence of at least one HRS state in both stored data states.

However, the read operation of the CRS cell is destructive for the memory state of "1" (LRS/HRS), because a positive read voltage greater than V_{th1} [Fig. 16(f)] triggers the filament in the bottom cell and places the CRS into the LRS/LRS or "ON" state. Since a measurable current is detected, the "1" memory state can be detected. After the read operation, a write-back process is required to return the device stack from the "ON" state back to memory state "1." This increases the complexity of the peripheral circuit design, and the extra power and time associated with the write-back operation significantly increases read power and degrades read bandwidth.

When the cell stores "0" (HRS/LRS), a positive readout voltage has no effect on the bottom cell, leading to low read current and no need for write-back. One challenge for empirical CRS device-pairs is attaining symmetric set and reset voltages in the two devices, lest the "ON" state become unstable.[13,102,103] By using an external series resistor,



(g) Possible states of the CRS.

CRS state	Memristive element A	Memristive element B	Resistance CRS
0	HRS	LRS	≈HRS
1	LRS	HRS	≈HRS
ON	LRS	LRS	LRS + LRS; only during read operation
OFF	HRS	HRS	≫HRS; only in initial state

State 'ON' occurs only during the read operation. State 'OFF' is the initial state without further importance for the operation and is present only after fabrication. The states '0' and '1' are the memory states for which the overall resistance of the CRS is the same: $R_{CRS} = R_{HRS} + R_{LRS} \approx R_{HRS}$ (with $R_{LRS} \ll R_{HRS}$).

FIG. 16. (Color online) By combining (a) bipolar resistive element A with a Pt/solid electrolyte/Cu stack with its (b) I-V-characteristic, together with (c) a reversed copy of the same cell, bipolar element B with a Cu/solid electrolyte/Pt stack and its (d) I-V-characteristic, the (e) stacked combination of elements A and B device acts as a CRS with a (f) access devicelike I-V-characteristic exhibiting (g) two stored data states amongst the four possible combinations of individual device states, both offering high resistance under half-select conditions. Reprinted with permission from Linn *et al.*, Nat. Mater. **9**, 403 (2010). Copyright 2010 Macmillan Publishers Ltd.

threshold voltages can be shifted to help obtain a stable "ON" state,[13,102] allowing the CRS concept to be applied to any bipolar resistive switching material, irrespective of the symmetry in set/reset voltages.[13] However, the use of external resistors further increases write power.

Rosezin *et al.* successfully demonstrated the vertical integration of CRS cells based on $Cu/SiO_2/Pt$ bipolar switches for passive crosspoint array applications with an OFF/ON resistance ratio of >1500 and switching speed of <120 μs.[104] Yu *et al.* have identified suitable pulse widths and pulse amplitudes for CRS devices based on Ag/GeSe elements.[13] In order to get similar read/write windows over a wide timescale, it is best that the set/reset switching dynamics of the two individual elements be the same, because the speed of the aggregate CRS device will be limited to the speed of the slowest switching process. A CRS device formed by antiserially connecting two RRAM elements with $Pt/ZrO_x/HfO_x/Pt$ structure in a crosspoint structure was reported by Lee *et al.*[105] In this demonstration, each ReRAM exhibited typical linear ON-current characteristics, yet acted as an access device for RESET-read-selection cycling of the other device.

Lee *et al.* reported a CRS device fabricated by connecting two $Pt/Ta_2O_{5-x}/TaO_{2-x}/Pt$ cells antiserially via external contacts as shown in Fig. 17.[106] Interestingly, the individual resistive memory cell in the mentioned stack had asymmetric I-V behavior [Figs. 17(a) and 17(b)] with the Schottky barrier at the Pt/Ta_2O_{5-x} interface helping to suppress leakage current for voltages between −0.7 and +0.7 V [Figs. 17(c) and 17(d)] by either being reverse biased [label (b) at far right side of Fig. 17(c)], or being below the threshold voltage of forward bias [label (a) at far right side of Fig. 17(c)].

In contrast to these CRS devices which require complex stacking of metal-oxide/metal layers, Nardi *et al.* recently demonstrated that CRS behavior can also be exhibited in single-layer nonpolar-RRAM devices,[107] by selectively severing the filament of oxygen vacancies either at the top or bottom interface. Then, instead of two separate devices each with its own filament, a single layer device either contains a filament severed at the top or at the bottom. This behavior was empirically observed (Fig. 18) in an oxide-RRAM device in simple $TiN/HfO_x/TiN$ structure with 5 nm thick HfO_x active layer.[107] In these devices, bipolar switching, CRS behavior, and unipolar switching could each be obtained by careful sequencing of switching operations and current compliance.[107] However, the current levels at which CRS behavior is obtained are fairly high, nor is it clear how the voltage and current conditions for CRS behavior might vary across a large array of such devices.

Yang *et al.* have reported CRS behavior in a tantalum oxide based single stack RRAM device,[108,109] attributed to the redistribution of oxygen vacancies in the tantalum oxide layers. As with the HfO_x work, systematic adjustment of the component materials and stoichiometries allows different switching behaviors—including unipolar, bipolar, and complementary switching (CS)—to be obtained in bilayer ($Pd/Ta_2O_{5-x}/TaO_y/Pd$) and trilayer stacks ($Pd/Ta_2O_{5-x}/TaO_y/TaO_z/Pd$).[109] In another study, CRS properties with low operating voltages (∼1 V) were reported in $Pt/TiO_{2-x}/TiN_xO_y/TiN$ by Tang *et al.*[110] One-step plasma oxidation of TiN film was used to partly oxidize the TiN bottom electrode, creating an oxygen reservoir layer of TiN_xO_y. The CRS behavior is attributed to redistribution of oxygen

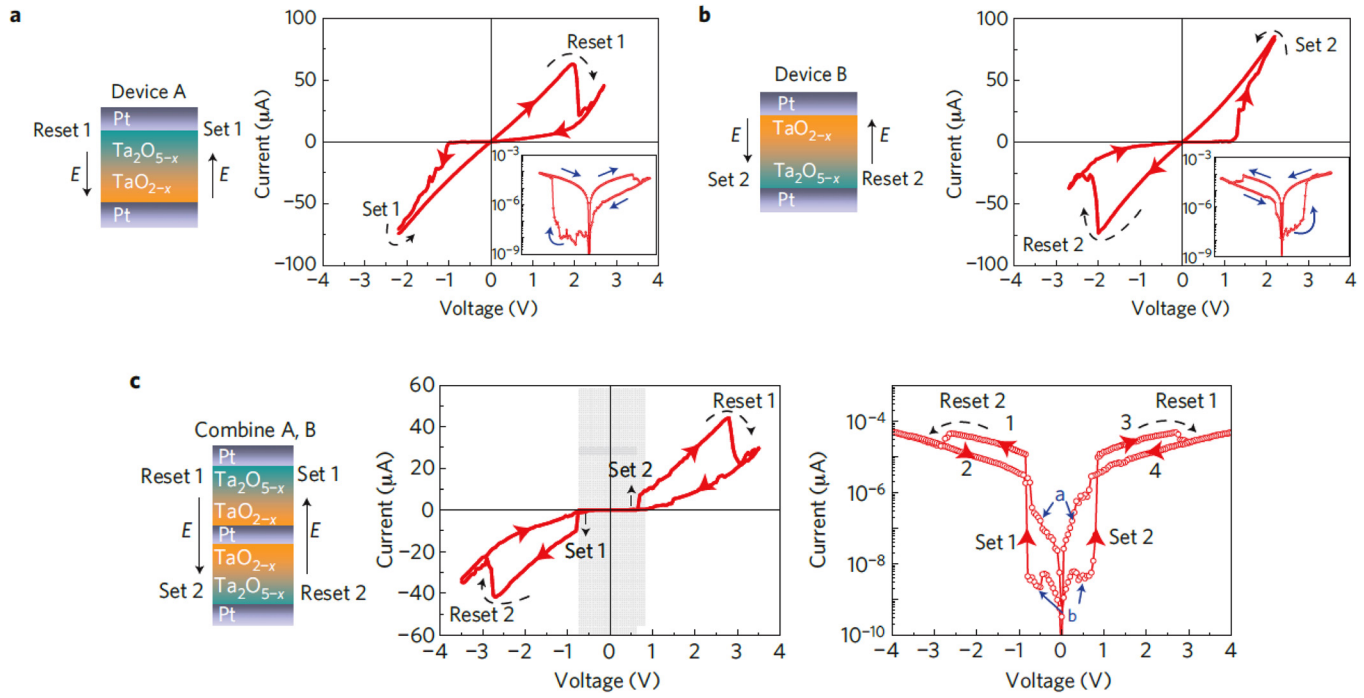


FIG. 17. (Color online) (a) Operation and schematic representation of a single resistive memory device. (b) A single reversed memory device. (c) A merged, antiserially connected device with a floating middle Pt electrode. Operation of the antiserial device shows the required switching region in between −0.7 and +0.7 V (gray region), which inhibits switching. Reprinted with permission from Lee *et al.*, Nat. Mater. **10**, 625 (2011). Copyright 2011 Macmillan Publishers Ltd.

vacancies between the Pt/TiO_{2-x} top interface and the TiO_{2-x}/TiN_xO_y bottom interface.

B. Hybrid RRAM-access devices

1. Hybrid devices using MIT

A hybrid memory is one in which a single device exhibits both memory as well as an access device properties. For instance, a combined metal–insulator transition and memory device was achieved by controlling oxygen concentration in a vanadium oxide film.[82] While stoichiometric VO_2 film exhibited typical MIT behavior with selection properties, and a nonstoichiometric V_2O_{5-x} film showed typical resistive memory switching behavior, vanadium oxide film with various intermediate oxygen content showed hybrid characteristics. Coexistence of multiple phases of vanadium oxide at the electrode interface region was observed by x-ray photoelectron spectroscopy (XPS).[82]

Similar to vanadium oxide but offering a significantly higher transition temperature, niobium oxide can exhibit both memory and access device properties for intermediate oxygen concentrations (Fig. 19).[84] In TEM and energy-dispersive X-ray spectroscopy (EDX) analysis, a 10 nm thick NbO_x layer is clearly observed. To confirm the high temperature stability of the hybrid memory device, DC measurement at 125 °C was implemented, showing uniform memory and access device characteristics even at high temperature conditions.[84]

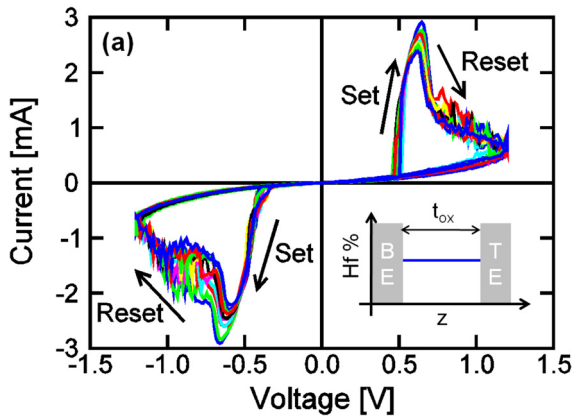


FIG. 18. (Color online) Measured I-V curves for symmetric RRAM, demonstrating CS in single stack nonpolar RRAM. Reprinted with permission from Nardi *et al.*, IEDM Tech. Dig. **2011**, 31.1. Copyright 2011 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

2. Hybrid devices using tunnel diodes

A scalable transistorless crosspoint technology was demonstrated by Meyer *et al.* by combining a novel oxide memory element and a cointegrated nonlinear tunnel diode.[111] The basic memory cell stack includes Pt as top and bottom electrodes, a thick (25 nm) crystalline perovskite conductive metal oxide (CMO) as the nonfilamentary switching layer, and a crystalline, high-quality dielectric tunnel oxide with the thickness of about 2–3 nm. Good crystallinity of both CMO and tunnel oxide are crucial in order to get desired memory and access device characteristics. Four layers of memory cells stacked in the BEOL above the silicon have been demonstrated in a 0.13 μm 64 Mb test chip.[112] BEOL integration is possible because the sputter deposition of both

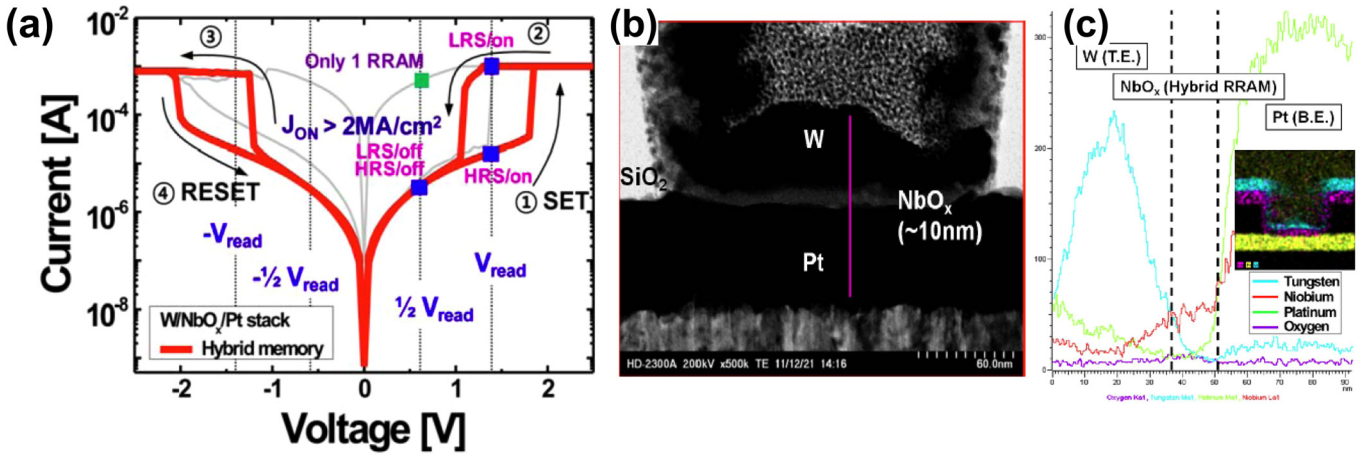


FIG. 19. (Color online) (a) I-V characteristics, (b) cross-section TEM image and (c) EDX line profile of hybrid (W/NbO_x/Pt) memory device. Reprinted with permission from S. Kim *et al.*, Symp. VLSI Technol. **2012**, T18.3. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

CMO and tunnel oxide materials can be performed at temperatures from 380 to 450 °C.[111]

The I-V curves of the aggregate memory device can be tuned by varying the tunnel oxide thickness, and fit well to a trap-assisted tunneling mechanism.[111] When positive voltage is applied to the top electrode (TE), the oxygen ions migrate toward it and excess negative charge accumulates in the tunnel oxide, which repels the tunneling electrons. As a result, barrier height increases, tunneling current decreases, and the measured device resistance increases [lower curves in Fig. 20(a)]. Conversely, negative voltage on TE forces the oxygen ions to move out of the tunnel oxide into the conductive metal oxide. Consequently, the tunneling current increases due to reduction in the effective barrier height, which leads to the low resistance state. The excellent scaling of initial, programmed and erased states with area [Fig. 20(b)] indicates the nonfilamentary switching mechanism.

No forming is required to obtain this resistive switching behavior. As both the resistance states are retained after program and erase operations, the memory device is nonvolatile with typical resistance ratio of ~10. A cycling endurance of $>10^6$ cycles with typical operation voltage of ±3 V and 10 μs pulse was obtained.[111] Since half-select leakage currents are fairly large, an asymmetric array favoring long wordlines and short bitlines is used.[112,113] The readout scheme is timed carefully to sense the device state before the steadily increasing leakage from half-selected lines overwhelms the signal.[112] Since read currents are fairly low, readout is slow; however, incorporation of local gain stages allows the read to occur in 5 μs (Ref. 114) instead of 50 μs.[113]

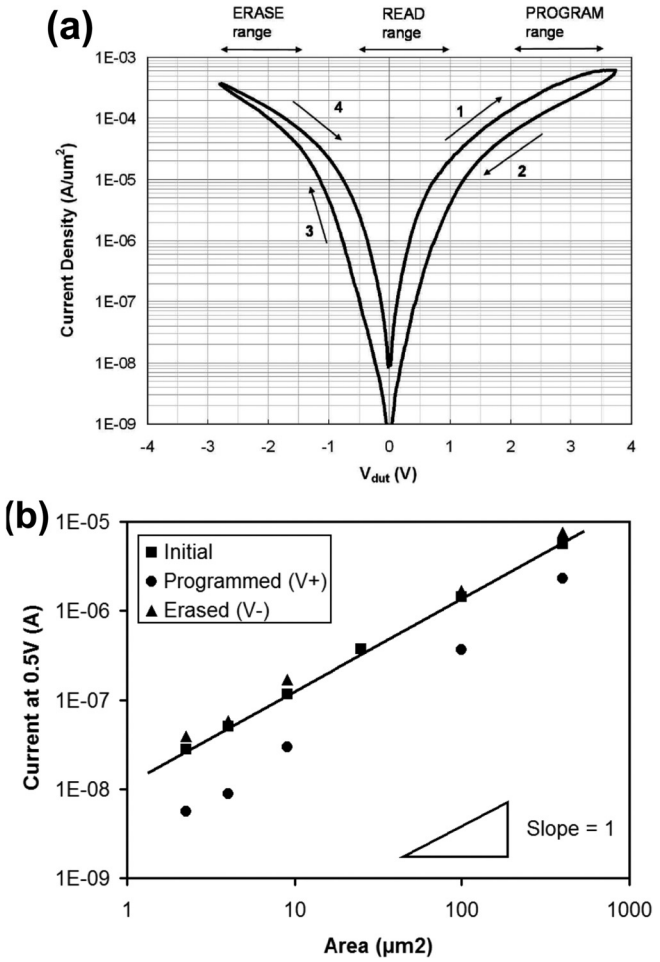


FIG. 20. (a) High field DC programming. Reprinted with permission from Chevallier *et al.*, *Proceedings of ISSCC* (2010), p. 14.3. Copyright 2010 IEEE. (b) Area scaling of initial, program and erase state. Reprinted with permission from Meyer *et al.*, Proceedings of NVMTS, 2008. Copyright 2008 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

C. Nonlinear RRAM

Another alternative for suppressing sneak-path leakage is to introduce a strong nonlinearity in the switching I-V curve of the memory element, by inserting a thin layer which acts as tunnel-barrier. Various types of nonlinear characteristics in RRAM device have been reported,[115–122] including a

2 Mb RRAM crosspoint array fabricated with 54 nm technology without access devices.[121]

Switching, which occurs at or near a TiO_x/Ta_2O_5 interface, can be attributed to either charge trapping/detrapping (type I switching, requiring high voltages) or oxygen vacancy migration (type II switching, requiring high switching currents). It has been reported that type II devices show better switching characteristics with higher switching speed (<10 ns), smaller operation voltage (<4 V), and better retention characteristics (>20 h at 150 °C).[121]

To achieve larger array size, low reset current (<3 μA) and high nonlinearity ($\gg$30) are needed, depending on driver capability.[121] Further improvement in I-V curves can be obtained either by using different spacer materials or by changing the Ti/O ratio. Park *et al*. have reported a nonlinear RRAM cell with ultralow operating current of <1 μA by engineering the switching oxide and by inserting thin barrier layers between oxide and electrodes. Such cells have been shown to switch with 1 μA current compliance and exhibit nonlinearities up to 17 [Fig. 21(c)].[122]

These two barrier layers have different roles in switching. In the bottom barrier, direct tunneling is dominant in the low voltage regime but Fowler–Nordheim tunneling takes place in the high voltage regime, leading to high nonlinearity in the LRS.[122] The top barrier provides self-current compliance, improving pulse endurance by suppressing current overshoot but also offering a wider operation margin for low current switching.[122] The overall conduction mechanism in LRS is Poole–Frenkel emission with tunneling through barrier layers shown in Figs. 21(a) and 21(b). Such devices can be switched between LRS and HRS states for $>10^7$ cycles with 1 μs pulses, while maintaining a resistance ratio of $\sim10^2$ and showing very little degradation in switching characteristics. As with conventional RRAM, multiple different LRS resistance levels can be obtained for MLC operation by controlling the current compliance.

The dependence of nonlinearity on the current compliance and thickness of the switching material was investigated by Lentz *et al*.[123] The device had a $TiN/TiO_2/Ti/Pt$ stack with two different TiO_2 thicknesses, 5 and 25 nm. For the thicker sample, the highest nonlinearity (of 7.4×) was achieved at the lowest Forming/SET currents, because as SET current increases, the thicker filament exhibits a more ohmic behavior. For the thinner (5 nm) sample, a higher nonlinearity (of 30×) was measured.[123]

Yang *et al*. successfully demonstrated a nonlinearity of $\sim$100 in TaO_x based devices by engineering the interface between TaO_x and TiO_{2-x} in a $Pt/TaO_x/TiO_{2-x}/Pt$ stack.[124] They found that the Pt/TaO_x interface was responsible for the resistive switching, while the oxide heterojunction of TaO_x/TiO_{2-x} with Schottky-like metal/semiconductor contact produced the nonlinearity. The operation of this device is attributed to electrothermal effects, and a new oxide phase is formed in the conducting channel where local

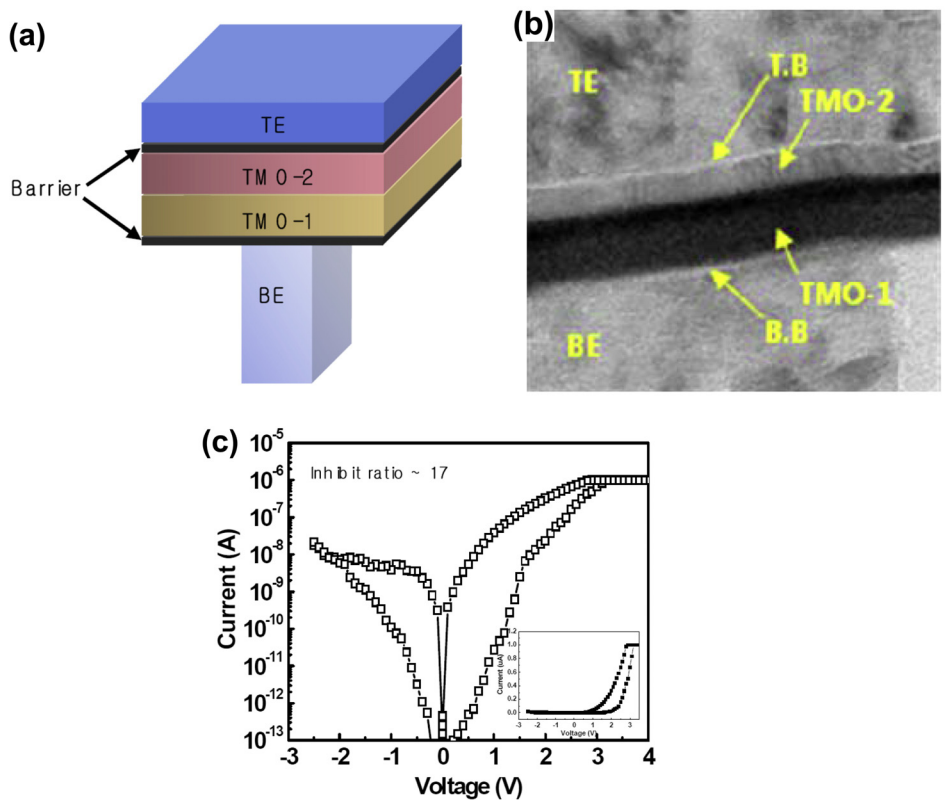


FIG. 21. (Color online) (a) Schematic of vertical RRAM structure. Thin barrier layer is inserted between TMO and electrode at both side. (b) X-TEM image ReRAM cell. (c) I-V curve of low current cell switching set current of $\sim$1 μA with nonlinear I-V characteristic. Reprinted with permission from Park *et al*., IEDM Tech. Dig. **2012**, 20.8. Copyright 2012 IEEE. Personal use of this material is permitted. However, permission to reuse this material for any other purpose must be obtained from the IEEE.

temperatures are higher because of Joule heating, leading to a negative differential resistance effect.

VI. CONCLUSION

In this review, we have seen that successful operation of large crosspoint arrays of "state-holding" memory devices requires the presence of strong nonlinearity at every crosspoint in order to achieve low-power programming and reliable memory readout. Any discrete access device explicitly added to provide this nonlinearity should offer a large ON-OFF ratio, from ultralow OFF leakage used when cells are in the half- and unselected states, to ultrahigh ON current densities for rapid switching of selected memory devices. Alternatively, this strong nonlinearity can be realized by the memory device itself. Fig. 22 shows a qualitative assessment of various stand-alone access devices. As discussed in the text and in Ref. 25, quantitative stand-alone assessment of devices is greatly complicated by the fact that the voltages at which the access device must deliver a high ON-OFF current contrast depend on the particular choice of NVM with which the access device is to be paired. For this reason, no quantitative table is shown in this paper.

After decades of research and development, Si-based access devices are very well understood. Three-terminal silicon devices are difficult to implement in a $4F^2$ footprint, while two-terminal devices based on single-crystal silicon (such as p-n diodes) preclude 3D multilayer stacking. Polysilicon devices can be stacked, but high ON-state current density can generally be achieved only with high ($>400\,^{\circ}C$) fabrication temperatures.

Access devices based on oxide PN junctions or on metal-oxide Schottky barriers offer both relative ease of integration and low ($<400\,^{\circ}C$) processing temperatures. However, the ON-state current density of both oxide PN junctions and metal-oxide Schottky barriers is several orders of magnitude smaller than what is required for operation with most resistive memory elements. In addition to current density limitations, both the Si rectifiers and the oxide PN junction diodes support only unipolar operation, and would not work with bipolar memory elements. However, it is possible to get bidirectional operation in metal-oxide-metal Schottky barrier devices by appropriate choice of metal electrodes.

Threshold switches constitute another category of potential access devices for nonvolatile memory devices. OTS devices, MIT–based devices, and the TVS are three potential candidates. OTS devices have been shown with good switching performance and in large arrays, but require complex materials and improvements in switching endurance and leakage current before being commercially viable. It is also not yet clear what range of NVM switching voltages and state resistances OTS or other threshold access devices might be able to support.

Selector	J_{ON}	Selectivity	Bidirectional	3D	Other challenges/questions/observations
Vertical Si transistor					Additional process complexity.
Si PN diode					Poly-Si pn diode may be suitable for 3D unipolar NVM.
Si Punchthrough diode					Scalability.
Oxide PN diode					
Oxide/nitride Schottky barriers					Relatively easy to integrate.
Varistor-type Bidirectional Switch (multilayer oxide barrier)					Unknown inter-device yield/variability; voltage margin for higher voltage NVM; Pt electrode processing; write endurance.
Chalcogenide threshold switch					Control of threshold voltage and its variability.
Insulator-metal transition switch					Transition temperature needs to be much higher than chip operating temperature.
Threshold Vacuum Switch					Unknown yield/variability; speed; manufacturability; effect of high-current pulse cycling on off-current.
MIEC selector					Voltage margin for higher voltage NVM.

FIG. 22. (Color online) Comparison table for stand-alone access devices (selectors). For J_{ON}, yellow (light gray) indicates the capability for passing maximum possible current densities >MA/cm^2, green (medium gray) is ~10 MA/cm^2, and red (dark gray) marks devices incapable of current densities beyond MA/cm^2. The *Selectivity* column describes the largest demonstrated ON-OFF ratio, with yellow (light gray) indicating $>10^4$ and green (medium gray) $>10^6$. By definition, the ON-OFF ratio available at a particular half-select voltage cannot be higher than this, and in general will be lower depending on selector threshold voltage, subthreshold slope, and the switching voltage of the companion NVM device. For *3*-D integration, the criteria for yellow (light gray) is a nonstandard (higher temperature) BEOL process, while green (medium gray) indicates full 400 °C BEOL compatibility.

In contrast, MIT and TVS switches have been demonstrated only at the few-device level, both individually and in series with prototype NVM devices. While MIT devices using materials such as NbO_2 with acceptably high threshold temperatures have been demonstrated, it will be important to further reduce the half-select and unselect leakage currents, which may be difficult given the narrow bandgaps of these materials. Also, higher threshold temperatures implies that added electrothermal power is being used to trigger the MIT behavior, which could potentially increase the overall power required for NVM switching. The TVS device is interesting, but it will be critical to demonstrate that sufficiently identical vacuum gaps, leading to tightly distributed electrical switching characteristics, can be demonstrated at high yield over large arrays of TVS access devices.

Oxide tunnel barriers offer steeply nonlinear curves, and have attained success in combination with various memories including CB-RAM and nonfilamentary CMO. However, it is extremely important that the switching voltages of the NVM remain low so that the half-select leakage—as evaluated at half the applied voltage across both memory and access device, as well as across any extra bias dissipated in the wiring—can still be low.

BEOL-friendly access devices based on copper-containing MIEC materials offer large currents ($>100\ \mu$A), bipolar operation, and ultralow leakage (<10 pA). Cointegration with PCM, integration in large (512 kbit) arrays with 100% yield and tight distributions, fast transient operation, long-term persistence of the required low-leakage, and scalability to aggressive technology nodes have all been demonstrated. However, larger voltage margin than the $V_m \sim 1.6$ V achieved to date will be required, if MIEC access devices are to be used with NVM devices with switching voltages larger than $\sim$1.3 V.

Methods for adding access device functionality to nonvolatile memory include the CRS, hybrid devices in which access device functionality is incorporated together with the memory functionality, and nonlinear RRAM, in which barrier layers introduce nonlinearity to help reduce leakage current. Although a CRS device can significantly reduce sneak-path currents, destructive readout with subsequent write-back is required. In addition, each CRS device must either be fabricated from a stack of two well-behaved and symmetric memory elements, or a single RRAM layer with very precise control over operating voltages. A particularly important step will be the demonstration of reliable CRS operation at the low switching currents (20–50 μA) required for implementation in the narrow-pitch wiring of advanced technology nodes.[21]

Another aspect of self-selected memories is the difficulty in independently tuning access device and memory functionality. This is more of an issue in hybrid devices using metal-insulator transitions or using tunnel barriers, which participate in the motion of oxygen ions, since the select function is nearly inseparable from the memory function. While nonlinearity can be tuned by increasing the thickness of an added tunnel barrier, the maximum nonlinearities achieved so far are well below the large (10^6–10^7) values needed for large (1000×1000) subarrays.

Most of the exciting applications for emerging nonvolatile memory technologies require that these devices are packed densely in vast arrays offering many gigabytes if not terabytes of solid-state storage. To reach this goal, the development of robust access device technologies—either as an explicit second device offering large nonlinearities, or as the engineering of an implicit strong nonlinearity into the state-bearing memory device—will likely be an important, if not critical, contributor.

[1]G. W. Burr, R. S. Shenoy, and H. Hwang, "Select device concepts for crossbar arrays," *Resistive Switching—From Fundamentals Of Nanoionic Redox Processes to Memristive Device Applications*, edited by D. Ielmini and R. Waser (Wiley–VCH, in press).
[2]E. P. G. Wright, Electric connecting device, U.S. patent 2,667,542 (26 January 1954).
[3]R. S. Williams, IEEE Spectr. **45**, 28 (2008).
[4]K. Kim and Y. J. Song, Integr. Ferroelectr. **61**, 3 (2004).
[5]S. Raoux *et al.*, IBM J. Res. Dev. **52**, 465 (2008).
[6]M. Hosomi *et al.*, IEDM Tech. Dig. **2005**, 459.
[7]*Resistive Switching—From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*, edited by D. Ielmini and R. Waser (Wiley–VCH, in press).
[8]R. Freitas and W. Wilcke, IBM J. Res. Dev. **52**, 439 (2008).
[9]G. W. Burr, B. N. Kurdi, J. C. Scott, C. H. Lam, K. Gopalakrishnan, and R. S. Shenoy, IBM J. Res. Dev. **52**, 449 (2008).
[10]A. Flocke and T. G. Noll, Eur. Solid State Circuits Conf. **2007**, 328.
[11]A. Flocke, T. G. Noll, C. Kugeler, C. Nauenheim, and R. Waser, 8th IEEE Conf. Nanotechnol. **2008**, 319.
[12]J. Liang and H.-S. P. Wong, IEEE Trans. Electron Devices **57**, 2531 (2010).
[13]S. Yu, J. Liang, Y. Wu, and H.-S. P. Wong, Nanotechnology **21**, 465202 (2010).
[14]A. Chen, *11th IEEE Conference on Nanotechnology*, Portland, OR, 15–18 August 2011, pp. 1767–1771.
[15]A. Chen and M.-R. Lin, IEEE Int. Reliab. Phys. Symp. **2011**, MY.7.1.
[16]C. Xu, X. Dong, N. P. Jouppi, and Y. Xie, *Design, Automation and Test in Europe Conference and Exhibition (DATE)* (2011), pp. 1–6.
[17]A. Chen, Z. Krivokapic, and M.-R. Lin, *70th Annual Device Research Conference (DRC)* (2012), pp. 219–220.
[18]J. Liang, S. Yeh, S. S. Wong, and H.-S. P. Wong, *2012 4th IEEE International Memory Workshop (IMW)* (2012), pp. 61–64.
[19]D. Niu, C. Xu, N. Muralimanohar, N. P. Jouppi, and Y. Xie, *Proceedings of the 2012 ACM/IEEE International Symposium on Low Power Electronics and Design* (2012), pp. 209–214.
[20]L. Zhang, S. Cosemans, D. J. Wouters, G. Groeseneken, and M. Jurczak *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC)* (2012), pp. 282–285.
[21]J. L. Liang, S. Yeh, S. S. Wong, and H. S. P. Wong, ACM J. Emerging Technol. Comput. Syst. **9** (2013).
[22]C. L. Lo, T. H. Hou, M. C. Chen, and J. J. Huang, IEEE Trans. Electron Devices **60**, 420 (2013).
[23]A. Chen, IEEE Trans. Electron Devices **60**, 1318 (2013).
[24]International Technology Roadmap for Semiconductors (2012). See http://www.itrs.net.
[25]P. Narayanan, G. W. Burr, R. S. Shenoy, S. Stephens, K. Virwani, A. Padilla, B. Kurdi, and K. Gopalakrishnan, *Device Research Conference*, (2014), V-A5.
[26]Y.-C. Chen, C. F. Chen, C. T. Chen, J. Y. Yu, S. Wu, S. L. Lung, R. Liu, and C.-Y. Lu, IEDM Tech. Dig. **2003**, 30.3.
[27]D. B. Strukov and K. K. Likharev, J. Nanosci. Nanotechnol. **7**, 4146 (2007).
[28]G. W. Burr *et al.*, Symp. VLSI Technol. **2012**, 41.
[29]A. Padilla *et al.*, IEDM Tech. Dig. **2010**, 29.4.
[30]A. Padilla *et al.*, J. Appl. Phys. **110**, 054501 (2011).
[31]H. S. Yoon *et al.*, Symp. VLSI Technol. **2009**, T2B-2.
[32]I. G. Baek *et al.*, IEDM Tech. Dig. **2011**, 31.8.
[33]H. Tanaka *et al.*, Symp. VLSI Technol. **2007**, 14.
[34]W. Kim, S. Choi, J. Sung, T. Lee, C. Park, H. Ko, J. Jung, I. Yoo, and Y. Park, Symp. VLSI Technol. **2009**, T10A-2.

[35]S. Whang *et al.*, IEDM Tech. Dig. **2010**, 29.7.
[36]X. P. Wang *et al.*, IEDM Tech. Dig. **2012**, 20.6.
[37]C.-H. Wang *et al.*, IEDM Tech. Dig. **2010**, 29.6.
[38]G. Servalli, IEDM Tech. Dig. **2009**, 5.7.1.
[39]J. H. Oh *et al.*, IEDM Tech. Dig. **2006**, 2.6.
[40]C. Zhang *et al.*, IEEE Electron Device Lett. **32**, 1014 (2011).
[41]K. S. Lee *et al.*, Jpn. J. Appl. Phys. **49**, 08JF03 (2010).
[42]H. An *et al.*, *2012 Proceedings of the European Solid–State Device Research Conference (ESSDERC)* (2012), pp. 149–152.
[43]K. S. Lee, J. J. Han, H. Lim, S. Nam, C. Chung, H. S. Jeong, H. Park, H. Jeong, and B. Choi, IEEE Electron Devices Lett. **33**, 242 (2012).
[44]Y. Sasago *et al.*, Symp. VLSI Technol. **2009**, T2B-1.
[45]S. B. Herner *et al.*, IEEE Electron Device Lett. **25**, 271 (2004).
[46]K. Huet, C. Boniface, R. Negru, and J. Venturini, AIP Conf. Proc. **1496**, 135 (2012).
[47]E.-X. Ping, Y. Erokhin, H.-J. Gossmann, and F. A. Khaja, *12th IEEE International Workshop on Junction Technology (IWJT)* (2012), pp. 24–28.
[48]Y. H. Song, S. Y. Park, J. M. Lee, H. J. Yang, and G. H. Kil, IEEE Electron Device Lett. **32**, 1023 (2011).
[49]G.-H. Kil, H.-J. Yang, G.-H. Lee, S.-H. Lee, and Y.-H. Song, Jpn. J. Appl. Phys., Part 1 **51**, 04DJ02 (2012).
[50]V. S. S. Srinivasan *et al.*, IEEE Electron Device Lett. **33**, 1396 (2012).
[51]S. Lashkare, P. Karkare, P. Bafna, M. V. S. Raju, V. S. S. Srinivasan, S. Lodha, U. Ganguly, J. Schulze, and S. Chopra, *Proceedings of IMW* (2013), pp. 178–181.
[52]M. H. Lee, C.-Y. Kao, C.-L. Yang, Y.-S. Chen, H. Y. Lee, F. Chen, and M.-J. Tsai, *69th Annual Device Research Conference (DRC)* (2011), pp. 89–90.
[53]Y.-S. Park, G.-H. Kil, and Y.-H. Song, Jpn. J. Appl. Phys., Part 1 **51**, 106501 (2012).
[54]I. G. Baek *et al.*, IEDM Tech. Dig. **2005**, 750.
[55]M. J. Lee *et al.*, Adv. Mater. **19**, 73 (2007).
[56]M. J. Lee *et al.*, IEDM Tech. Dig. **2007**, 30.2.
[57]B. S. Kang *et al.*, Adv. Mater. **20**, 3066 (2008).
[58]S.-E. Ahn, B. S. Kang, K. H. Kim, M.-J. Lee, C. B. Lee, G. Stefanovich, C. J. Kim, and Y. Park, IEEE Electron Devices Lett. **30**, 550 (2009).
[59]G. Tallarida, N. Huby, B. Kutrzeba-Kotowska, S. Spiga, M. Arcari, G. Csaba, P. Lugli, A. Redaelli, and R. Bez, Proceeding of International Memory Workshop, IMW'09, 2009.
[60]J.-J. Huang, C.-W. Kuo, W.-C. Chang, and T.-H. Hou, Appl. Phys. Lett. **96**, 262901 (2010).
[61]W. Y. Park, G. H. Kim, J. Y. Seok, K. M. Kim, S. J. Song, M. H. Lee, and C. S. Hwang, Nanotechnology **21**, 195201 (2010).
[62]G. H. Kim *et al.*, Appl. Phys. Lett. **100**, 213508 (2012).
[63]J.-J. Huang, Y.-M. Tseng, W.-C. Luo, C.-W. Hsu, and T.-H. Hou, IEDM Tech. Dig. **2011**, 31.7.
[64]J. J. Huang, Y. M. Tseng, C. W. Hsu, and T. H. Hou, IEEE Electron Device Lett. **32**, 1427 (2011).
[65]S. R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968).
[66]D. Ielmini, Phys. Rev. B **78**, 035308 (2008).
[67]D. Ielmini and Y. G. Zhang, J. Appl. Phys. **102**, 054517 (2007).
[68]D. Adler, M. S. Shur, M. Silver, and S. R. Ovshinsky, J. Appl. Phys. **51**, 3289 (1980).
[69]M. Anbarasu, M. Wimmer, G. Bruns, M. Salinga, and M. Wuttig, Appl. Phys. Lett. **100**, 143505 (2012).
[70]I. V. Karpov, M. Mitra, D. Kau, G. Spadini, Y. A. Kryukov, and V. G. Karpov, Appl. Phys. Lett. **92**, 173501 (2008).
[71]M. Wuttig and N. Yamada, Nat. Mater. **6**, 824 (2007).
[72]D. Krebs, S. Raoux, C. T. Rettner, G. W. Burr, M. Salinga, and M. Wuttig, Appl. Phys. Lett. **95**, 082101 (2009).
[73]W. Czubatyj and S. J. Hudgens, Electron. Mater. Lett. **8**, 157 (2012).
[74]A. C. Warren, IEEE Trans. Electron Devices **20**, 123 (1973).
[75]A. Pirovano, A. L. Lacaita, A. Benvenuti, F. Pellizzer, and R. Bez, IEEE Trans. Electron Devices **51**, 452 (2004).
[76]D. Emin, Phys. Rev. B **74**, 035206 (2006).
[77]I. V. Karpov, M. Mitra, D. Kau, G. Spadini, Y. A. Kryukov, and V. G. Karpov, J. Appl. Phys. **102**, 124503 (2007).
[78]D. C. Kau *et al.*, IEDM Tech. Dig. **2009**, 27.1.
[79]M.-J. Lee *et al.*, IEDM Tech. Dig. **2012**, 2.6.
[80]M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).
[81]M. Son *et al.*, IEEE Electron Device Lett. **32**, 1579 (2011).
[82]M. Son *et al.*, IEEE Electron Device Lett. **33**, 718 (2012).
[83]F. A. Chudnovskii, L. L. Odynets, A. L. Pergament, and G. B. Stefanovich, J. Solid State Chem. **122**, 95 (1996).
[84]S. Kim *et al.*, Symp. VLSI Technol. **2012**, T18.3.
[85]X. Liu, S. M. Sadaf, M. Son, J. Shin, J. Park, J. Lee, S. Park, and H. Hwang, Nanotechnology **22**, 475702 (2011).
[86]C. H. Ho *et al.*, IEDM Tech. Dig. **2012**, 2.8.
[87]A. Fazio, MRS Bull. **29**, 814 (2004).
[88]E. Cimpoiasu, S. K. Tolpygo, X. Liu, N. Simonian, J. E. Kukens, R. F. Klie, Y. Zhu, and K. K. Likharev, J. Appl. Phys. **96**, 1088 (2004).
[89]K. K. Likharev, Appl. Phys. Lett. **73**, 2137 (1998).
[90]W. Lee *et al.*, Symp. VLSI Technol. **2012**, T5.2.
[91]J. Woo *et al.*, Symp. VLSI Technol. **2013**, T12.4.
[92]A. Kawahara *et al.*, IEEE J. Solid-State Circuits **48**, 178 (2013).
[93]J. Shin *et al.*, J. Appl. Phys. **109**, 033712 (2011).
[94]S. Kim, W. Lee, and H. Hwang, *13th International Workshop on Cellular Nanoscale Networks and Their Applications (CNNA)* (2012), pp. 1–2.
[95]W. Lee *et al.*, ACS Nano **6**, 8166 (2012).
[96]J. Woo, D. Lee, E. Cha, S. Lee, S. Park, and H. Hwang, Appl. Phys. Lett. **103**, 202113 (2013).
[97]K. Gopalakrishnan *et al.*, Symp. VLSI Technol. **2010**, 19.4.
[98]R. S. Shenoy *et al.*, Symp. VLSI Technol. **2011**, T5B-1.
[99]K. Virwani *et al.*, IEDM Tech. Dig. **2012**, 2.7.
[100]G. W. Burr *et al.*, Symp. VLSI Technol. **2013**, T6.4.
[101]A. Padilla *et al.*, Device Res. Conf. **2014**, III-53.
[102]E. Linn, R. Rosezin, C. Kuegeler, and R. Waser, Nat. Mater. **9**, 403 (2010).
[103]D. J. Wouters *et al.*, IEEE Electron Devices Lett. **33**, 1186 (2012).
[104]R. Rosezin, E. Linn, L. Nielen, C. Kugeler, R. Bruchhaus, and R. Waser, IEEE Electron Device Lett. **32**, 191 (2011).
[105]J. Lee *et al.*, IEDM Tech. Dig. **2010**, 19.5.
[106]M.-J. Lee *et al.*, Nat. Mater. **10**, 625 (2011).
[107]F. Nardi, S. Balatti, S. Larentis, and D. Ielmini, IEDM Tech. Dig. **2011**, 31.1.
[108]Y. Yang, P. Sheridan, and W. Lu, Appl. Phys. Lett. **100**, 203112 (2012).
[109]Y. Yang, S. H. Choi, and W. Lu, Nano Lett. **13**, 2908 (2013).
[110]G. Tang, F. Zeng, C. Chen, H. Liu, S. Gao, C. Song, Y. Lin, G. Chen, and F. Pan, Nanoscale **5**, 422 (2013).
[111]R. Meyer, L. Schloss, J. Brewer, R. Lambertson, W. Kinney, J. Sanchez, and D. Rinerson, Proceedings of NVMTS, November 2008.
[112]C. J. Chevallier, C. H. Siau, S. F. Lim, S. R. Namala, M. Matsuoka, B. L. Bateman, and D. Rinerson, *Proceedings of ISSCC* (2010), p. 14.3.
[113]C. Chevallier, *Proceedings of 2011 Leti Workshop on Innovative Memory Technologies* (2011), p. 15.
[114]C. Chevallier, Proceedings of ITRS workshop on Storage Class Memory, July 2012.
[115]S. H. Jo and W. Lu, Nano Lett. **8**, 392 (2008).
[116]K.-H. Kim, S. H. Jo, S. Gaba, and W. Lu, Appl. Phys. Lett. **96**, 053106 (2010).
[117]Q. Zuo *et al.* IEEE Electron Devices Lett. **31**, 344 (2010).
[118]M. Yang, P. Hu, J. Q. Lu, Q. B. Lv, and S. W. Li, Appl. Phys. Lett. **98**, 213501 (2011).
[119]X. A. Tran *et al.*, IEDM Tech. Dig. **2011**, 31.2.
[120]Y. L. Huang, R. Huang, Y. Pan, L. J. Zhang, Y. M. Cai, G. Y. Yang, and Y. Y. Wang, IEEE Trans. Electron Devices **59**, 2277 (2012).
[121]H. D. Lee *et al.*, Symp. VLSI Technol. **2012**, T18.1.
[122]S.-G. Park *et al.*, IEDM Tech. Dig. **2012**, 20.8.
[123]F. Lentz, B. Roesgen, V. Rana, D. J. Wouters, and R. Waser, IEEE Electron Device Lett. **34**, 996 (2013).
[124]J. J. Yang *et al.*, Appl. Phys. Lett. **100**, 113501 (2012).