# An Electronic Synapse Device Based on Metal Oxide Resistive Switching Memory for Neuromorphic Computation

Shimeng Yu, *Student Member, IEEE*, Yi Wu, Rakesh Jeyasingh, Duygu Kuzum, and H.-S. Philip Wong, *Fellow, IEEE*

***Abstract*—The multilevel capability of metal oxide resistive switching memory was explored for the potential use as a single-element electronic synapse device. TiN/$HfO_x$/$AlO_x$/Pt resistive switching cells were fabricated. Multilevel resistance states were obtained by varying the programming voltage amplitudes during the pulse cycling. The cell conductance could be continuously increased or decreased from cycle to cycle, and about $10^5$ endurance cycles were obtained. Nominal energy consumption per operation is in the subpicojoule range with a maximum measured value of 6 pJ. This low energy consumption is attractive for the large-scale hardware implementation of neuromorphic computing and brain simulation. The property of gradual resistance change by pulse amplitudes was exploited to demonstrate the spike-timing-dependent plasticity learning rule, suggesting that metal oxide memory can potentially be used as an electronic synapse device for the emerging neuromorphic computation system.**

***Index Terms*—Bio-inspired system, neuromorphic computation, resistive switching memory, spike-timing-dependent plasticity (STDP), synapse.**

## I. Introduction

In THE LAST decades, the rapid progress in digital electronic system based on complementary metal–oxide–semiconductor (CMOS) integrated circuit technology has successfully opened up an information age. However, learning and intelligence have not yet been captured in the classical von Neumann computation paradigm. Energy efficiency is one of the major bottlenecks for the today's digital electronic system to be competitive in processing the information and data in a complex real-world environment. To perform a cortical simulation at the complexity of the cat brain even at a rate of a hundred times slower than the real-time neuron firing rate, the IBM's Blue Gene supercomputer demands heavy computation resources, requiring clusters of 147 456 microprocessors and 144 TB of memories, which consume a power of 1.4 MW [1], [2]. However, the power dissipation in the more complex human brain is only on the order of 10 W [3]. Thus, creating cognitive intelligent and energy-efficient systems is a grand challenge [4]. Neuromorphic computation is such a bio-inspired paradigm that could be more energy efficient than the conventional Boolean logic computation due to the parallelism and would have attractive features such as fault tolerance and capacity for adaptive learning [5]–[8]. The ability to simulate the functions of a brain will lead to a deeper understanding of neuroscience in a way that complements the physiological measurements of neural systems.

A typical neuromorphic computation system comprises neuron circuits and synapses. The synapses are the connections between the neuron circuits. Each neuron may have more than 1000 synapse connections with other neurons. Therefore, the major challenge for the hardware implementation of neuromorphic computation system is to develop the electronic synapse element that has high density and low energy consumption. The number of biological synapses is enormous ($\sim 10^{15}$) in the human brain [3]. As an order-of-magnitude estimation, the synapses operate at about a frequency of 1$\sim$10 Hz [3], and the human brain consumes a power on the order of 10 W [3]. Thus, on the average, the energy consumption per synaptic event is around 1$\sim$10 fJ. Therefore, the requirements of the electronic synapse include not only the ultrahigh integration density but also ultralow energy consumption. Furthermore, for an electronic synapse element to emulate the functions of the biological synapse, e.g., to be capable for the adaptive learning in a neural network, the synapse devices should exhibit plasticity in gradual change of the conductance. Combined with the proper design of the neuron circuits, the synapse devices should enable the functionality of Hebbian's learning rules, such as the spike-timing-dependent plasticity (STDP) [9].

Electronic synapse was first made by the circuits with tens of CMOS transistors and capacitors [10], which occupied significant area and consume substantial energy even if fabricated with the modern nanoscale CMOS technology. Efforts have been made to use a single three-terminal transistor as the electronic synapse, including a floating gate transistor [11], a ferroelectric transistor [12], and an organic nanoparticle transistor [13], for which the threshold voltage can be gradually changed by the amount of trapped charge, ferroelectric dipole polarization, or charged nanoparticles, respectively. Recently,

Manuscript received March 1, 2011; revised April 13, 2011; accepted April 17, 2011. Date of publication June 9, 2011; date of current version July 22, 2011. This work was supported in part by DARPA SyNAPSE (DSO Program Manager T. Hylton), by the National Science Foundation (NSF) under NSF ECCS 0950305, by the Nanoelectronics Research Initiative of the Semiconductor Research Corporation through the NSF/NRI Supplement to the NSF NSEC Center for Probing the Nanoscale, and by the member companies of the Stanford Non-Volatile Memory Technology Research Initiative. The work of S. Yu was supported by the Stanford Graduate Fellowship. The work of Y. Wu was supported by a fellowship from the O. G. Villard Engineering Fund at Stanford. The review of this paper was arranged by Editor H. Shang.

The authors are with the Center for Integrated Systems and Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: simonyu@stanford.edu; hspwong@stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2011.2147791

two-terminal solid-state memories that show electrically triggered resistive switching phenomenon have been proposed to serve as the electronic synapse devices [14], [15] with a potential for a 3-D integrated architecture. A polymeric electrochemical element [16], a memory capacitor based on a field-configurable ion-doped polymer [17], a Ag/a-Si programmable metallization cell [18], a $GdO_x/Cu:MoO_x$-based resistive switching device [19], and $TiO_2/TiO_{2-x}$-based resistive switching device [20] have been demonstrated for the electronic synapse devices. In this paper, we present a novel electronic synapse candidate using the $HfO_x/AlO_x$-based metal oxide resistive switching memory with attractive features such as sufficient resistance modulation window ($>3$ orders of magnitude), reasonable endurance cycles ($\sim 10^5$), low programming voltage ($<3$ V), fast programming speed ($\sim$10 ns), and low energy consumption per operation (subpicojoule). The property of low energy consumption is shown for the first time in two-terminal solid-state memory-based electronic synapse devices.

Metal oxide resistive switching memory has been regarded as a promising candidate for future nonvolatile memory application due to the simple structure, the scalability to the nanometer regime, and the compatibility with silicon CMOS technology [21], [22]. The mechanism of the resistive switching phenomenon in metal oxides has been widely attributed to the formation/rupture of nanoscale conductive filaments (CFs) which may consist of oxygen vacancies or metal precipitates [23], [24]. The set process, switching from high-resistance state (HRS) to low-resistance state (LRS), is attributed to the dielectric soft breakdown and creation of the CFs paths in the oxide matrix, while the reset process, switching from LRS to HRS, is attributed to the annihilation of the CFs by the Joule heating dissolution or the electrochemical reactions involving oxygen ions/vacancies [25], [26]. Among the metal oxide materials, $HfO_x$-based devices have shown excellent performance such as fast switching speed, excellent switching endurance, and reliable data retention [27]. In addition, 4-Mb $HfO_x$ memory arrays were successfully demonstrated [28]. At the same time, $AlO_x$-based devices have shown low switching energy consumption [29]. Moreover, the $HfO_x$-based memory with embedded $AlO_x$ buffer layer has shown better switching uniformity due to the incorporation of Al atoms into $HfO_x$ matrix materials [30]. Here, we investigate the use of $HfO_x/AlO_x$ bilayer structured devices as a two-terminal electronic synapse, focusing on the gradual resistance change capability by control of the input pulse amplitudes, the measurement of energy consumption per operation and the potential to exhibit the STDP learning rule [9] by proper design of the spike signaling scheme.

## II. Experimental Details

Synapse devices with $TiN/HfO_x/AlO_x/Pt$ thin-film stacks were fabricated. TiN is used as the top electrode, the oxide matrix consists of $HfO_x/AlO_x$, and Pt is used as the bottom electrode. A 50-nm Pt was first deposited by e-beam evaporation on silicon substrate as the bottom electrode layer. Then, a 5-nm $AlO_x$ was deposited by atomic layer deposition (ALD) using TDMA-Al (tetrakis dimethylamido aluminum $Al[N(CH_3)_2]_4$) and $H_2O$ as precursors at 300 °C, and then, a 5-nm $HfO_x$ was deposited by ALD using TEMA-Hf (tetrakis ethylmethylamino hafnium $Hf[N(C_2H_5)(CH_3)]_4$) and $H_2O$ as precursor at 220 °C. The crossbar patterns with $0.5 \times 0.5$ $\mu m^2$ active cell area were defined by photolithography. Then, a 50-nm TiN was deposited by reactive sputtering and was lifted off as the top electrode layer. The electrical measurements were performed with the Agilent 4156C semiconductor parameter analyzer, the Agilent 81101A pulse generator, and the Tektronix DPO 4054 oscilloscope. The bottom electrode was grounded, and the signals were applied to the top electrode in all the measurements.

## III. Results and Discussion

Fig. 1(a) illustrates an analogy between the biological synapse and the electronic synapse based on the metal oxide resistive switching memory. The biological synapse changes its conductance because the preneuron releases $Ca^{2+}$ or $Na^{2+}$ ions when the input impulses are received; similarly, the electronic synapse changes its conductance due to the oxygen vacancies generation and the oxygen ions migration when the input impulses are received. Fig. 1(b) shows the $I$–$V$ characteristics of the electronic synapse device measured by the typical dc double sweep. During the positive bias ramp, the stop voltage is 2 V, and a 100-$\mu$A compliance current is enforced to limit the current through the device. The device resistance can be set from the HRS to the LRS. During the negative bias ramp, the stop voltage is $-3.3$ V and no compliance current is enforced. The device resistance can be reset from LRS to HRS. Increase (set process) and decrease (reset process) in device conductance is equivalent to the potentiation and depression, respectively, in biological synapses. The devices show bipolar switching behavior, which means the set process and reset process occur at opposite bias polarities, which is attributed to the ability of the TiN top electrode to serve as an oxygen reservoir [32]. The resistance ratio between HRS and LRS is more than 1000, and the distinguished binary states can be used to store logic "0" or "1" for memory application. To emulate the functions of a biological synapse, however, we wish to obtain an analog memory which has multiple states in between HRS and LRS. Fig. 1(c) shows the current–voltage ($I$–$V$) characteristics of the device obtained measured by a modified dc double sweep. The sweep sequence is indicated by the number in the figure: first, we perform a continuous set (potentiation) with a consecutive increase of the set compliance from 1 to 200 $\mu$A, and the resistance of the device gradually decreases from HRS to LRS. Then, we perform a continuous reset (depression) with a consecutive increase of the magnitude of the reset stop voltage from $-1.3$ to $-3.3$ V, and the resistance of the device gradually increases from LRS to HRS. It is speculated that the gradual set process is due to the stronger CFs with lager diameters or multiple CFs formed with a larger current through the oxide matrix [33]; and the gradual reset is due to the partial annihilation of the CFs with a larger ruptured length of the CFs or a fewer number of CFs [34]. Although the method of using the set compliance current or the reset stop voltage in a dc sweep to modulate the LRS/HRS resistances has been reported before [27], here, the feature that the resistance can be

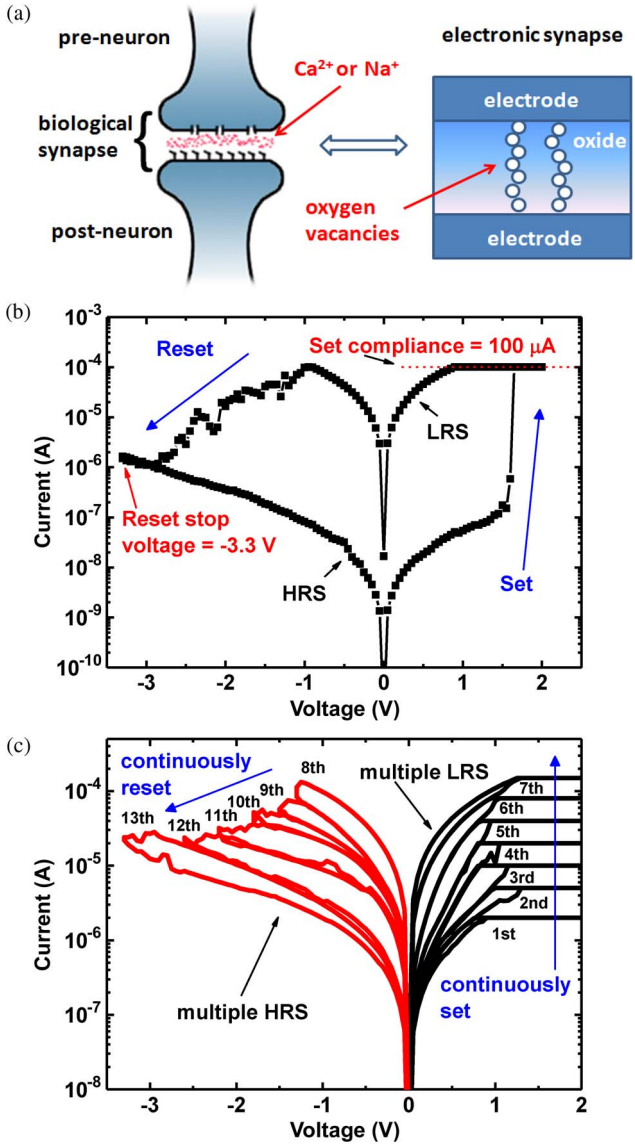



Fig. 1. (a) Analogy between the biological synapse and the electronic synapse based on the metal oxide resistive switching memory. The electronic synapse has a metal/insulator/metal thin-film structure. (b) $I$–$V$ characteristics of the $HfO_x/AlO_x$ synapse device measured by a typical dc double sweep. During the positive bias ramp, the stop voltage is 2 V, and 100 $\mu$A compliance current is enforced; during the negative bias ramp, the stop voltage is $-3.3$ V, and no compliance current is enforced. Binary states were obtained. (c) $I$–$V$ characteristics of the $HfO_x/AlO_x$ synapse device measured by a modified dc double sweep. The sweep sequence is indicated by the number in the figure: first, perform a continuous set (potentiation) with a consecutive increase of the set compliance from 1 to 200 $\mu$A, and then, perform a continuous reset (depression) with a consecutive increase of the reset stop voltage from $-1.3$ to $-3.3$ V. The resistances are continuously changed, and multiple states were obtained.

continuously increased or decreased between the intermediate states without going back to the original state is crucial for electronic synapse application.

The practical application of electronic synapse devices are under pulse signal inputs rather than the voltage ramp in a dc sweep. Therefore, we investigated the property of gradual resistance change in the context of pulse cycling. For the pulse programming, no compliance current was enforced. The elimination of the use of a current limiter is possible due to the fact that the pulse width is very short ($\sim$50 ns); thus, the excess damage to the cell during the set process is much reduced as compared to the dc sweep case. Fig. 2(a) shows the resistance evolution of the device for the first 100 pulse cycles. The pulsewidths were fixed to be 50 ns. First, starting from the LRS, the gradual reset process was performed; five pulses with the amplitudes consecutively increased from $-2.2$ to $-2.6$ V were applied, and the resistance was gradually increased from around 10 k$\Omega$ to several hundreds of kiloohms. Then, the gradual set process was performed; five pulses with the amplitudes consecutively increased from 1.4 to 1.8 V were applied, and the resistance was gradually decreased from several hundreds of kiloohms to around 10 k$\Omega$. After this 10-pulse sequence, the cell returned to the original LRS. The up–down evolution could be repeated for many times. It is speculated that a higher set pulse amplitude would impose more stress on the oxide materials to create more defects, e.g., oxygen vacancies, and more defects would lead to a lower resistance because the defect density would determine the local conductivity of the CFs [34]. Similarly, a higher reset pulse amplitude would recover more defects, and fewer remaining defects would lead to a higher resistance. It is also noted that there are noticeable resistance variations in the evolution, which is inevitable due to the stochastic nature of the resistive switching: the percolation property of the thin dielectric soft breakdown [35]. Fig. 2(b) shows the pulse cycling endurance data for up to $10^5$ cycles. Pulses of $\pm 3$ V with 50-ns width were applied to switch the device between HRS and LRS, and the resistances were read out by a 100-mV voltage at each sampling point. The HRS resistance tends to decrease because the unrecovered defects in the oxide matrix accumulate from cycle to cycle. However, the resistance window is still remarkable, up to $10^5$ cycles. Fig. 2(c) shows the similar up–down evolution as Fig. 2(a) but after the device underwent the $10^5$-cycle endurance test. Although the highest resistance value tends to decrease a bit, which reflects the degradation of the HRS in Fig. 2(b), the device still retains the property of the gradual resistance change. It is noted that most of the cells in our devices fail into a very low resistance state, around $10^6$ cycles probably due to too many unrecoverable oxygen vacancies. This is not an excellent value for synapse application (ideally, $> 10^9$ cycles is desired assuming spikes arrive at 10 Hz for ten years) but a reasonable value considering our simple fabrication conditions. Using a more refined process, $HfO_x$-based memory devices have shown endurance up to $10^{10}$ cycles by others [27]. Fig. 3 shows the stability of the device for different resistance states measured at an elevated temperature (100 °C). The resistance states are monitored by a 0.1-V read voltage every 10 s, and no significant change of the states is observed for a period of 7200 s (2 h). In general, the retention requirement for synapse devices is not as strict as the nonvolatile memory, because the synapse would receive the signals occasionally from the neuron anyway, and those signals would refresh its state. The aforementioned results suggest the $HfO_x/AlO_x$-based synapse devices have promising reliability for potential practical applications.

Another key concern for a practical electronic synapse device is the energy consumption per operation. Fig. 4 shows the observed waveforms of the applied programming voltage

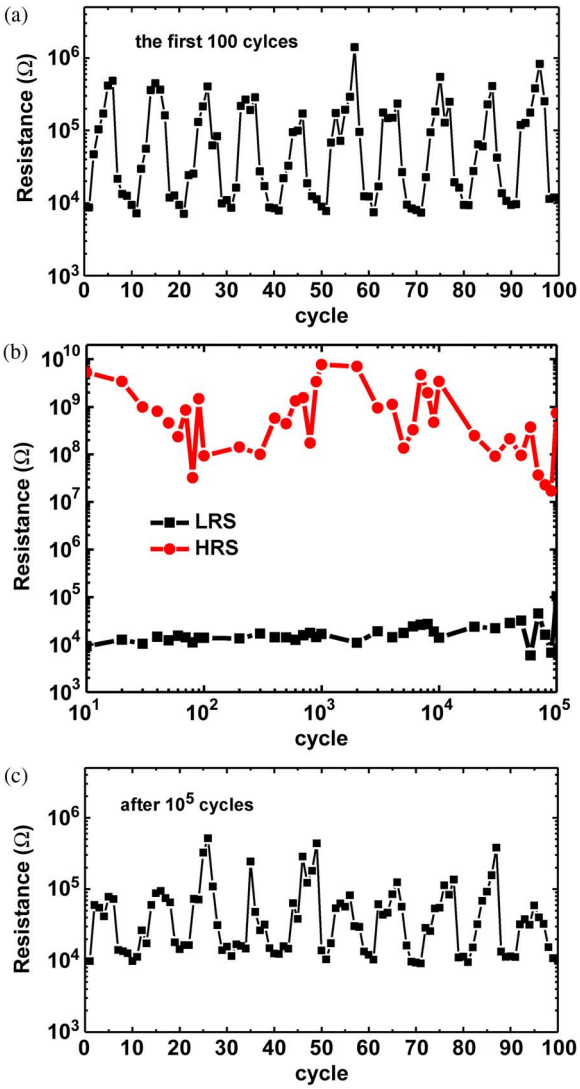



Fig. 2. (a) Resistance evolution of the $HfO_x/AlO_x$ synapse device for the first 100 pulse cycles. The pulsewidths were fixed to be 50 ns. First, starting from the LRS, the gradual reset process was performed: five pulses with the amplitudes consecutively increased from $-2.2$ to $-2.6$ V with a step of $-0.1$ V were applied. Then, the gradual set process was performed: five pulses with the amplitudes consecutively increased from 1.4 to 1.8 V with a step of 0.1 V were applied. (b) Endurance test to $10^5$ cycles. Pulses of $\pm 3$ V with 50-ns width were applied to switch the device between the HRS and the LRS, and the resistances were read out by a 100-mV voltage at each sampling point. (c) Resistance evolution of the same measurement as (a) performed on the same device but after the $10^5$-cycle endurance test.

pulses and the current response measured through a 50-Ω series resistor. In Fig. 4(a), the cell is set to be the lowest resistance state ($\sim 10\,\mathrm{k}\Omega$) before the reset operation, and then, a $-2.5$ V/50 ns reset pulse is applied. The current drops after a wait time around 10 ns, indicating that the reset occurs at that time. Then, the cell resistance is read to be around 1 MΩ. In Fig. 4(b), a $+1.5$ V/50 ns set pulse is applied. The current rises after a wait time around 30 ns, indicating that the set occurs at that time. By integrating the voltage, current, and time from these waveforms, we obtain the energy consumption per operation: 5.9 pJ for set and 6 pJ for reset. The example in Fig. 4 corresponds to the maximum energy consumption per operation because, here, we operate the cell in the lowest resistance

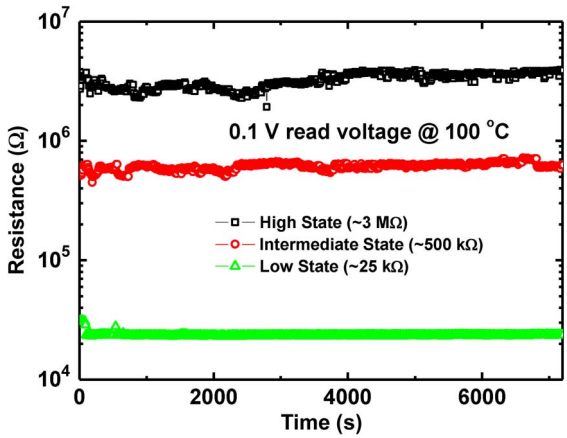



Fig. 3. Stability property for different resistance states measured at an elevated temperature (100 °C). The resistance states are monitored by a 0.1-V read voltage every 10 s, and no significant change of the states are observed for a time period of 7200 s (2 h).

state ($\sim 10\,\mathrm{k}\Omega$), and the electronic synapse's nominal operation regime is the intermediate resistance states ($\sim 300\,\mathrm{k}\Omega$). Thus, the nominal energy consumption per operation would fall into the subpicojoule range because the current through the cell is inversely proportional to the resistance value as a first-order approximation. As estimated before, the energy consumption per synaptic event in the human brain is in the 1–10-fJ range, and our electronic synapse devices roughly meet this requirement within an order of magnitude.

Next, we investigate the capability of electronic synapses for emulating the adaptive learning rules of the biological synapses, of which one of the most fundamental ones is the STDP learning rule [9]. A typical electronic neuron circuit and its synapse connections are shown in Fig. 5. The neuron circuit has adders at the input to sum all the signals from other neuron circuits, and the signals are accumulated at the leaky integrator, e.g., a capacitor. When the voltage at the leaky integrator reaches the threshold, this neuron fires, and it generates voltage spikes transmitted to other neurons. Furthermore, different neurons fires independently, and the spikes arrive at a specific synapse occasionally from both preneuron and postneuron in the opposite direction. The conductance of the synapse would change according to these spikes. Implementing the STDP learning rule requires that the changes of the synapse conductance to be a function of the time elapsed between the arrival of the prespike and that of the postspike [36]. The biological synapse data [36] is shown in Fig. 6. If the prespike precedes the postspike, the strength of the synapse undergoes a long-term potentiation (LTP) (increase in conductance, indicating that the connection between two neurons becomes stronger). Otherwise, it undergoes a long-term depression (LTD) (decrease in conductance, indicating that the connection between two neurons becomes weaker). Moreover, closer spike timing leads to a larger change of the conductance (Fig. 6). The relative change is defined as $\Delta G = (G_{\mathrm{after}} - G_{\mathrm{before}})/G_{\mathrm{before}}$. $G_{\mathrm{before}}$ is the conductance before the pre and post spike pair, and $G_{\mathrm{after}}$ is the conductance after the prespike and postspike pair. According to this definition, the range of $\Delta G$ for potentiation is $(0, +\infty)$, and for depression, it is $(-1, 0)$.

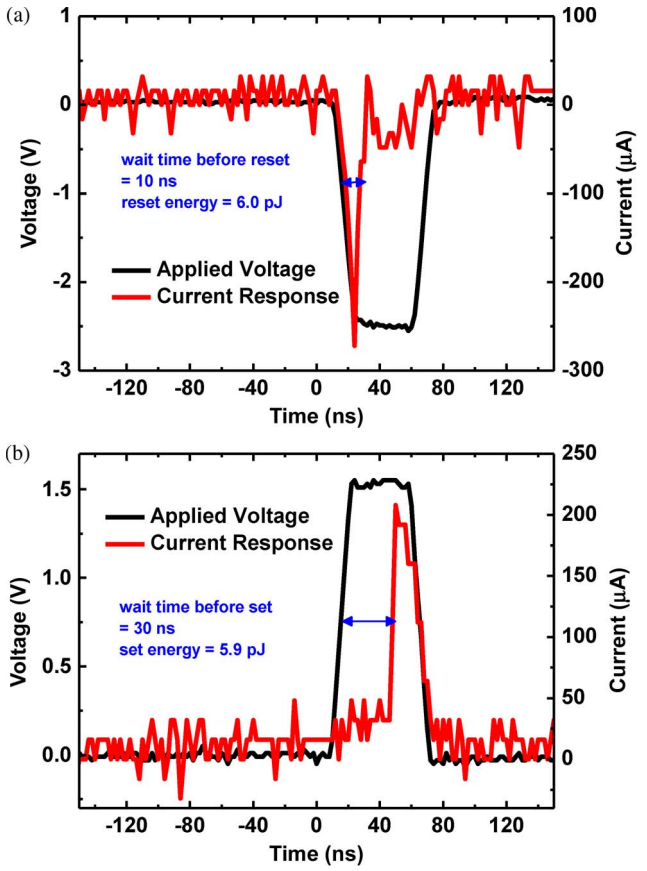



Fig. 4. (a) Applied voltage pulse and transient current response for a reset operation starting from the lowest state ($\sim$10 k$\Omega$). The pulse amplitude, width, rising time, and falling time is set to be $-2.5$ V, 50 ns, 10 ns, and 10 ns, respectively. (b) Applied voltage pulse and transient current response for a set operation starting from the high state ($\sim$1 M$\Omega$). The pulse amplitude, width, rising time, and falling time is set to be $+1.5$ V, 50 ns, 10 ns, and 10 ns, respectively.

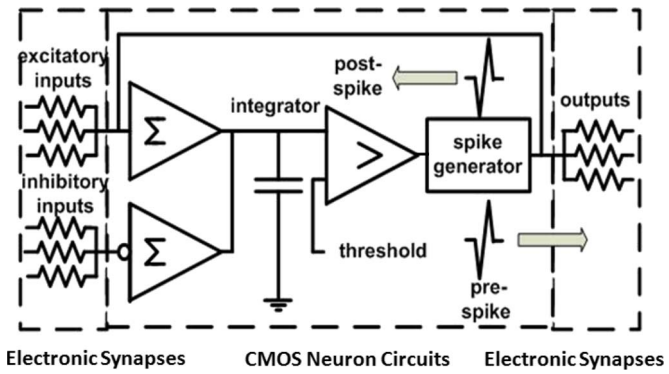



Fig. 5. Neuron circuit with electronic synapse devices. The neuron circuit has adders at the input to sum all the signals from other neuron circuits, and the signals are accumulated at the leaky integrator, e.g., a capacitor. When the voltage at the leaky integrator reaches the threshold, this neuron fires and it generates voltage spikes transmitted to other neurons. The synapse's conductance would change according to the relative arrival times of the prespike and postspike.

If the resistance of an electronic synapse device can be tuned by the input pulse amplitude or pulsewidth, a time-division multiplexing (TDM) approach [37] can be used to design the shape of the spike to demonstrate the STDP behavior. Signals through the electronic synapses are constructed by multiplexing

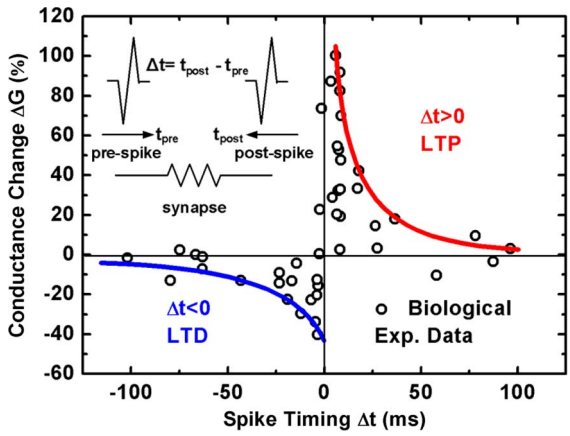



Fig. 6. STDP data of a biological synapse adapted from [36]. If the prespike precedes the postspike, the synapse's conductance undergoes an LTP (increase in conductance). Otherwise, it undergoes an LTD (decrease in conductance). Closer spike timing leads to larger relative change of the conductance. The relative change is defined as $\Delta G = (G_{after} - G_{before})/G_{before}$, where $G_{before}$ is the conductance before the prespike and postspike pair, and $G_{after}$ is the conductance after the prespike and postspike pair. According to this definition, the range of $\Delta G$ for potentiation is $(0, +\infty)$, and for depression, it is $(-1, 0)$.

the timeslots in the time domain. The key idea is to use TDM to translate the difference of spiking timing into the difference of pulsewidth. In [18], the authors used a signaling scheme that tuned the pulsewidth in each timeslot because the resistance of their devices could be modulated by the pulsewidth. Owing to the exponential voltage–time relationship [38] in the switching dynamics, it was found that linearly varying the pulse amplitudes can tune the resistance as effectively as exponentially varying the pulsewidth but with considerably less energy consumption [39]. To make the programming energy-efficient for synapse devices, in this paper, we developed a signaling scheme that tunes the pulse amplitude in each time slot [40]. Furthermore, the generation of these specific signals is done by the corresponding neuron circuit design in practical application.

Let us first examine how the pulse amplitude in a single pulse can effectively tune the resistance states. Fig. 7(a) shows the resistance modulation of five states achieved by varying the amplitudes of positive potentiation pulses from a very high initial state ($>$ 1 M$\Omega$). Fig. 7(b) shows the resistance modulation of five states achieved by varying the amplitudes of negative depression pulses from a very low initial state ($\sim$10 k$\Omega$). The aforementioned two initial states correspond to two extreme conditions, and in general, the initial state is an intermediate state that can undergo either potentiation or depression. Therefore, we investigated this case in more detail. Fig. 8(a) shows a resistance modulation of ten states achieved by varying the amplitudes of positive potentiation pulses and negative depression pulses from an intermediate initial state (200 k$\Omega$ $\sim$ 300 k$\Omega$, i.e., the shaded region in the plot). As we can see, for a linear increase of positive pulse amplitude from 1.6 to 2 V, the resistance gradually decreases to around 30 k$\Omega$. For a linear increase of the negative pulse amplitude from $-2.4$ to $-2.8$ V, the resistance gradually increases to around 3 M$\Omega$. To exploit the capability of the pulse amplitude modulation for STDP, the shape of the spike is designed as in Fig. 8(b). The

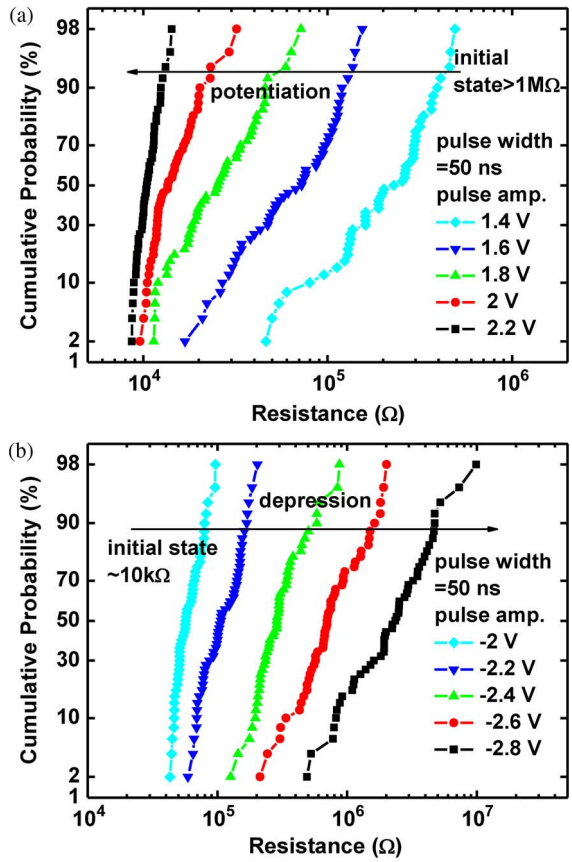



Fig. 7. (a) $HfO_x/AlO_x$ synapse device resistance modulation (potentiation part) by the pulse amplitude from an initial state with a very high resistance value ($> 1$ M$\Omega$). These measurements were performed by applying a single pulse (50 ns) with various amplitudes (1.4 to 2.2 V) and then recording the resistance of the final state. (b) $HfO_x/AlO_x$ synapse device resistance modulation (depression part) by the pulse amplitude starting from an initial state with a very high resistance value ($\sim 10$ k$\Omega$). These measurements were performed by applying a single pulse (50 ns) with various amplitudes ($-2$ to $-2.8$ V) and then recording the resistance of the final state.

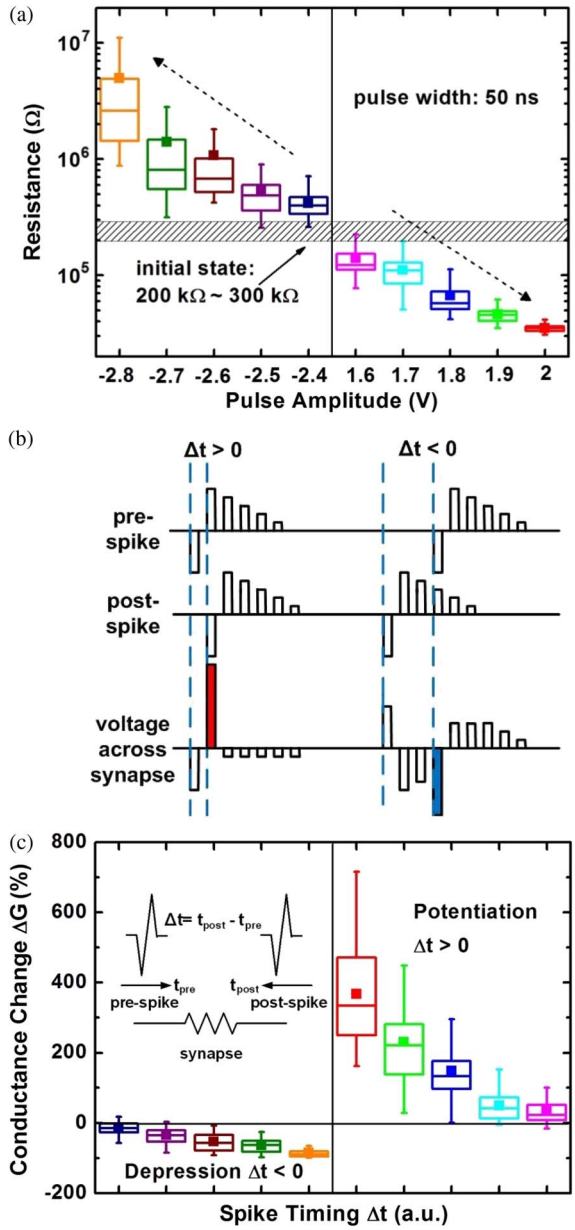



Fig. 8. (a) The $HfO_x/AlO_x$ synapse device resistance modulation (both potentiation and depression part) by the pulse amplitude starting from an initial state with an intermediate resistance value (200 k$\Omega \sim$ 300 k$\Omega$, the shaded region in the plot). These measurements were performed by applying a single pulse (50 ns) with various amplitudes (the potentiation pulses varied from 1.6 to 2V, and the depression pulses varied from $-2.4$ to $-2.8$ V) and then recording the resistance of the final state. Each error bar consists of 50 independent tests. (b) STDP realization schemes developed with TDM and pulse amplitude modulation. The pulse amplitudes for the prespike are $-1.4$, 1, 0.9, 0.8, 0.7, and 0.6 V, consecutively, and for the postspike, they are $-1$, 1.4, 1.3, 1.2, 1.1, and 1 V, consecutively. (c) STDP-like curve calculated with the data (a) employing the signal schemes in (b).

spike consists of a series of single pulses in consecutive timeslots. A negative pulse occupies the first timeslot; then, positive pulses with decreasing amplitudes follow in subsequent timeslots. The design principle is that the no single pulse in any timeslot can affect the device's resistance. Only when there is an overlap of the prespike and postspike can a programming pulse be generated for which the amplitude is large enough to modulate the resistance. The voltage dropped on the synapse is defined as the voltage of the prespike minus the voltage of the postspike. Thus, the negative pulse in the first timeslot acts as an enable function to determine which positive pulse could contribute to the programming pulse. Closer spike timing means the negative pulse would overlap a positive pulse with larger amplitude, leading to a larger resistance change. If the prespike precedes the postspike, a positive programming pulse (the red one in the figure) is produced. Otherwise, a negative one (the blue one in the figure) is produced. An example design of the spike is outlined here. The period for each pulse in one spike is designed to be 1 $\mu$s. The duty cycle for each positive pulse can be very short, e.g., 50 ns, since the actual programming event happens very fast, as shown in Fig. 4. However, the duty cycle for the first negative pulse should be relatively longer, e.g., 500 ns, since it acts as an enable function that must tolerate the timing variability caused by the $RC$ delay through the interconnect. Therefore, the total time for one spike is 6 $\mu$s, corresponding to a maximum operating frequency is 167 kHz. Biological synapses asynchronously receive spikes from neurons at frequencies of about 1–10 Hz [3]. Therefore, the total time required for one spike in our programming scheme is not a limitation, indicating that the neuromophic

system can work at a higher frequency than the biological system. Using the signaling scheme explained earlier, the STDP-like curve could be reproduced in our electronic synapse devices as Fig. 8(c). The relative conductance changes were calculated with the data in Fig. 8(a). Owing to the fast switching speed of our devices, the actual programming time in each pulse is designed to be very short (approximately tens of nanoseconds) to minimize the energy consumption. However, the timing in the signal scheme has an arbitrary unit because it can be flexibly designed by the length of the timeslot to fit the typical STDP time window (approximately in milliseconds) in biological synapses. It should be noted that the STDP curve in biological synapses shown in Fig. 6 is only one of many STDP behaviors observed in experimental measurements. There are different forms of STDP-like behaviors that can enable adaptive learning as illustrated in [41]. In this paper, the electronic synaptic element (the metal oxide resistive switching device) changes its conductance based on the signaling scheme [Fig. 8(b)]. The signaling scheme is flexible so that it can be modified according to the form of STDP-like behavior desired. The signaling scheme is implemented by the neuron circuits rather than the device itself. This way, a variety of STDP behavior can be emulated on the same chip using the same electronic synapse device technology. The aforementioned results suggest that the $HfO_x/AlO_x$-based memory devices have the potential to be trained and achieve learning for the adaptive computation in neuromorphic systems.

## IV. Conclusion

Device performance such as the low energy consumption per operation (subpicojoule), reasonable endurance cycles ($\sim 10^5$), and the capability to modulate the resistance based on the input pulse amplitudes suggest that the $HfO_x/AlO_x$-based memory is a competitive candidate for the electronic synapse devices application. The number of state levels needed depends on the specific system-level algorithm. At the device level, the capability to offer multilevel states can allow more flexibility for the system-level design. The present challenges are the notable variation of the switching phenomenon in metal oxide resistive switching memory. Because the potentiation and depression of the biological synapses or learning in neural systems is not deterministic, some degree of variation will be tolerated. Thus, the requirement for variations in a neuromorphic system is usually less strict than that for a multibit data storage system, because it relies on self-healing networks that implement massively parallel computations [14]. For example, instead of a single pulse, the average effect on hundreds of parallel synapse inputs into one neuron (see the circuit diagram in Fig. 5) determines whether the neuron will fire or not. Therefore, there is no need to completely eliminate the randomness of the resistive switching due to the percolation nature of the CF generation. However, control of the randomness to some degree is necessary. Recent research work of scaling down the metal oxide memory device's cell area to sub-50-nm regime [42] shows a remarkable improvement of switching uniformity. Thus, further reducing the device area to the nanometer regime may enable the metal oxide resistive switching memory to be the ultrahigh-dense ultralow-energy- consumption reliable electronic synapse device.

## Acknowledgment

The authors would like to thank Dr. Y. Chai, Dr. J. Provine, and Cambridge Nanotech for the device fabrication, Dr. S.-B. Kim for the measurement setup, and Prof. K. Boahen for the discussion.

## References

[1] R. Ananthanarayanan, S. K. Esser, H. D. Simon, and D. S. Modha, "The cat is out of the bag: Cortical simulations with $10^9$ neurons and $10^{13}$ synapses," in *Proc. IEEE/ACM Conf. High Perform. Netw. Comput.*, 2009, pp. 1–12.

[2] D. Adee (2009). IBM unveils a new brain simulator. *IEEE Spectr.* [Online]. Available: http://spectrum.ieee.org/computing/hardware/ibm-unveils-a-new-brain-simulator

[3] E. R. Kandel and J. H. Schwartz, *Principles of Neural Science*, 2nd ed. New York: Elsevier, 1985.

[4] Interdisciplinary challenges beyond the scaling limits of Moore's law.

[5] C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, no. 10, pp. 1629–1636, Oct. 1990.

[6] S. Fusi, M. Annunziato, D. Badoni, A. Salamon, and D. Amit, "Spike-driven synaptic plasticity: Theory, simulation, VLSI implementation," *Neural Comput.*, vol. 12, no. 10, pp. 2227–2258, Oct. 2000.

[7] J. V. Arthur and K. Boahen, "Learning in silicon: Timing is everything," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2006, pp. 75–82.

[8] K. K. Likharev, "Hybrid CMOS/nanoelectronic circuits: Opportunities and challenges," *J. Nanoelectron. Optoelectron.*, vol. 3, no. 3, pp. 203–230, Dec. 2008.

[9] S. Song, K. D. Miller, and L. F. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nat. Neurosci.*, vol. 3, no. 9, pp. 919–926, Sep. 2000.

[10] G. Indiveri, E. Chicca, and R. Douglas, "A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 211–221, Jan. 2006.

[11] H. Kosaka, T. Shibata, H. Ishii, and T. Ohmi, "An excellent weight-updating-linearity EEPROM synapse memory cell for self-learning neuron-MOS neural networks," *IEEE Trans. Electron Devices*, vol. 42, no. 1, pp. 135–143, Jan. 1995.

[12] S. M. Yoon, E. Tokumitsu, and H. Ishiwara, "An electrically modifiable synapse array composed of metal-ferroelectric-semiconductor (MFS) FETs using $SrBi_2Ta_2O_9$ thin films," *IEEE Electron Device Lett.*, vol. 20, no. 5, pp. 229–231, May 1999.

[13] F. Alibart, S. Pleutin, D. Guerin, C. Novembre, S. Lenfant, K. Lmimouni, C. Gamrat, and D. Vuillaume, "An organic nanoparticle transistor behaving as a biological spiking synapse," *Adv. Funct. Mater.*, vol. 20, no. 2, pp. 330–337, Jan. 2010.

[14] G. S. Snider, "Self-organized computation with unreliable, memristive nanodevices," *Nanotechnology*, vol. 18, no. 36, p. 365 202, Sep. 2007.

[15] J. Borghetti, Z. Li, J. Straznicky, X. Li, D. A. A. Ohlberg, W. Wu, D. R. Stewart, and R. S. Williams, "A hybrid nanomemristor/transistor logic circuit capable of self-programming," *Proc. Nat. Acad. Sci.*, vol. 106, no. 6, pp. 1699–1703, Feb. 2009.

[16] A. Smerieri, T. Berzina, V. Erokhin, and M. P. Fontana, "Polymeric electrochemical element for adaptive networks: Pulse mode," *J. Appl. Phys.*, vol. 104, no. 11, p. 114 513, Dec. 2008.

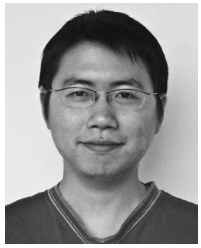
[17] Q. Lai, L. Zhang, Z. Li, W. F. Stickle, R. S. Williams, and Y. Chen, "Analog memory capacitor based on field-configurable ion-doped polymers," *Appl. Phys. Lett.*, vol. 95, no. 21, p. 213 503, Nov. 2009.

[18] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano Lett.*, vol. 10, no. 4, pp. 1297–1301, Apr. 2010.

[19] H. Choi, H. Jung, J. Lee, J. Yoon, J. Park, D.-J. Seong, W. Lee, M. Hasan, G.-Y. Jung, and H. Hwang, "An electrically modifiable synapse array of resistive switching memory," *Nanotechnology*, vol. 20, no. 34, p. 345 201, Aug. 2009.

[20] K. Seo, I. Kim, S. Park, S. Jung, M. Jung, J. Park, J. Kong, K. Lee, B. Lee, and H. Hwang, "Resistive switching device for neuromorphic device application," in *Proc. Int. Conf. Solid State Devices Mater.*, 2010.

[21] R. Waser, R. Dittmann, G. Staikov, and K. Szot, "Redox-based resistive switching memories—Nanoionic mechanisms, prospects, and challenges," *Adv. Mater.*, vol. 21, no. 25/26, pp. 2632–2663, Jul. 2009.
[22] S. Yu, B. Lee, and H.-S. P. Wong, "Metal oxide resistive switching memory," in *Functional Metal Oxide Nanostructures*, J. Q. Wu, Ed. Berlin, Germany: Springer-Verlag, 2011, to be published.
[23] M.-J. Lee, S. Han, S. H. Jeon, B. H. Park, B. S. Kang, S.-E. Ahn, K. H. Kim, C. B. Lee, C. J. Kim, I.-K. Yoo, D. H. Seo, X.-S. Li, J.-B. Park, J.-H. Lee, and Y. Park, "Electrical manipulation of nanofilaments in transition-metal oxides for resistance-based memory," *Nano Lett.*, vol. 9, no. 4, pp. 1476–1481, Apr. 2009.
[24] D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim, and C. S. Hwang, "Atomic structure of conducting nanofilaments in $TiO_2$ resistive switching memory," *Nat. Nanotechnol.*, vol. 5, no. 2, pp. 148–153, Feb. 2010.
[25] U. Russo, D. Ielmini, C. Cagli, and A. L. Lacaita, "Filament conduction and reset mechanism in NiO-based resistive-switching memory (RRAM) devices," *IEEE Trans. Electron Devices*, vol. 56, no. 2, pp. 186–192, Feb. 2009.
[26] D. S. Jeong, H. Schroeder, and R. Waser, "Mechanism for bipolar switching in a Pt/$TiO_2$/Pt resistive switching cell," *Phys. Rev. B, Condens. Matter*, vol. 79, no. 19, p. 195 317, May 2009.
[27] H. Y. Lee, Y. S. Chen, P. S. Chen, P. Y. Gu, Y. Y. Hsu, S. M. Wang, W. H. Liu, C. H. Tsai, S. S. Sheu, P. C. Chiang, W. P. Lin, C. H. Lin, W. S. Chen, F. T. Chen, C. H. Lien, and M.-J. Tsai, "Evidence and solution of over-RESET problem for $HfO_x$ based resistive memory with sub-ns switching speed and high endurance," in *IEDM Tech. Dig.*, 2010, pp. 460–463.
[28] S.-S. Sheu, M.-F. Chang, K.-F. Lin, C.-W. Wu, Y.-S. Chen, P.-F. Chiu, C.-C. Kuo, Y.-S. Yang, P.-C. Chiang, W.-P. Lin, C.-H. Lin, H.-Y. Lee, P.-Y. Gu, S.-M. Wang, F. T. Chen, K.-L. Su, C.-H. Lien, K.-H. Cheng, H.-T. Wu, T.-K. Ku, M.-J. Kao, and M.-J. Tsai, "4 Mb embedded SLC resistive-RAM macro with 7.2 ns read-write random-access time and 160 ns MLC-access capability," in *Proc. IEEE ISSCC Tech. Dig.*, 2011, pp. 200–202.
[29] Y. Wu, B. Lee, and H.-S. P. Wong, "$Al_2O_3$-based RRAM using atomic layer deposition (ALD) with 1 $\mu$A reset current," *IEEE Electron Device Lett.*, vol. 31, no. 12, pp. 1449–1451, Dec. 2010.
[30] S. Yu, B. Gao, H. B. Dai, B. Sun, L. F. Liu, X. Y. Liu, R. Q. Han, J. F. Kang, and B. Yu, "Improved uniformity of resistive switching behaviors in $HfO_2$ thin films with embedded Al layers," *Electrochem. Solid-State Lett.*, vol. 13, no. 2, pp. H36–H38, 2010.
[31] M. Fujimoto, H. Koyama, M. Konagai, Y. Hosoi, K. Ishihara, S. Ohnishi, and N. Awaya, "$TiO_2$ anatase nanolayer on TiN thin film exhibiting high-speed bipolar resistive switching," *Appl. Phys. Lett.*, vol. 89, no. 22, p. 223 509, Nov. 2006.
[32] Y. Sato, K. Tsunoda, K. Kinoshita, H. Noshiro, M. Aoki, and Y. Sugiyama, "Sub-100-$\mu$A reset current of nickel oxide resistive memory through control of filamentary conductance by current limit of MOSFET," *IEEE Trans. Electron Devices*, vol. 55, no. 5, pp. 1185–1191, May 2008.
[33] L. Goux, Y.-Y. Chen, L. Pantisano, X.-P. Wang, G. Groeseneken, M. Jurczak, and D. J. Wouters, "On the gradual unipolar and bipolar resistive switching of TiN\$HfO_2$\Pt memory systems," *Electrochem. Solid-State Lett.*, vol. 13, no. 6, pp. G54–G56, 2010.
[34] M. J. Rozenberg, M. J. Sánchez, R. Weht, C. Acha, F. Gomez-Marlasca, and P. Levy, "Mechanism for bipolar resistive switching in transition-metal oxides," *Phys. Rev. B, Condens. Matter*, vol. 81, no. 11, p. 115 101, Mar. 2010.
[35] S. Blonkowski, "Filamentary model of dielectric breakdown," *J. Appl. Phys.*, vol. 107, no. 8, p. 084 109, Apr. 2010.
[36] G.-Q. Bi and M.-M. Poo, "Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type," *J. Neurosci.*, vol. 18, no. 24, pp. 10 464–10 472, Dec. 1998.
[37] G. S. Snider, "Spike-timing-dependent learning in memristive nanodevices," in *Proc ACM/IEEE NANOARCH*, 2008, pp. 85–92.
[38] S. Yu and H.-S. P. Wong, "A phenomenological model for the reset mechanism of metal oxide RRAM," *IEEE Electron Device Lett.*, vol. 31, no. 12, pp. 1455–1457, Dec. 2010.
[39] S. Yu, Y. Wu, and H.-S. P. Wong, "Investigating the switching dynamics and multilevel capability of bipolar metal oxide resistive switching memory," *Appl. Phys. Lett.*, vol. 98, no. 10, p. 103 514, Mar. 2011.
[40] S. Yu and H.-S. P. Wong, "Modeling the switching dynamics of programmable-metallization cell (PMC) memory and its application as synapse device for a neuromorphic computation system," in *IEDM Tech. Dig.*, 2010, pp. 520–523.
[41] H. Z. Shouval, S. S.-H. Wang, and G. M. Wittenberg, "Spike timing dependent plasticity: A consequence of more fundamental learning rules," *Front. Comput. Neurosci.*, vol. 4, pp. 19-1–19-13, 2010.
[42] J. Lee, J. Shin, D. Lee, W. Lee, S. Jung, M. Jo, J. Park, K. P. Biju, S. Kim, S. Park, and H. Hwang, "Diode-less nano-scale $ZrO_x$/$HfO_x$ RRAM device with excellent switching uniformity and reliability for high-density cross-point memory applications," in *IEDM Tech. Dig.*, 2010, pp. 452–455.



**Shimeng Yu** (S'10) received the B.S. degree from Peking University, Beijing, China, in 2009, and the M.S. degree from Stanford University, Stanford, CA, in 2011, where he is currently working toward the Ph.D. degree with the Department of Electrical Engineering.

He is the author or a coauthor of one book chapter in Functional Metal Oxide Nanostructures, and more than 30 papers on different topics, including several papers in IEDM, Symposia of VLSI Technology, IEEE TRANSACTIONS ON ELECTRON DEVICES, IEEE ELECTRON DEVICE LETTERS, *Applied Physics Letters*, *Nanotechnology*, etc. He also serves as an active reviewer for *Applied Physics Letters*, *Nanotechnology*, *Journal of Physics D: Applied Physics*, *Journal of the Electrochemical Society*, and *Electrochemical and Solid-State Letters*. His past research actives includes the simulation of the parameters fluctuation in nanoscale transistors and SRAM cells. He is currently working on the fabrication, characterization, and modeling of the emerging resistive switching memory devices and their applications for neuromorphic computation systems.

Mr. Yu received the Stanford Graduate Fellowship from 2009 to 2011 and an IEEE Electron Devices Society Masters Student Fellowship in 2010.

**Yi Wu** received the B.S. degree from Peking University, Beijing, China, in 2008, and the M.S. degree from Stanford University, Stanford, CA, in 2010, where she is currently working toward the Ph.D. degree with the Department of Electrical Engineering.

She was a Research Intern with Industrial Technology Research Institute, Hsinchu, Taiwan, in 2010. She is currently working on metal-oxide-based resistance switching memory with a focus on the thin-film property study and the memory structure innovations.

**Rakesh Jeyasingh** received the B.E. degree in electronics and communication engineering from Anna University, Chennai, India, in 2005 and the M.E. degree in microelectronics from the Indian Institute of Science, Bangalore, India, in 2007. He is currently working toward the Ph.D. degree with the Department of Electrical Engineering, Stanford University, Stanford, CA, since 2008.

His research interests include nonvolatile memory design and modeling with specific emphasis to phase change memories. His past experience includes custom digital circuit design and high-speed system design.

**Duygu Kuzum** received the B.S. degree in electrical engineering from Bilkent University, Ankara, Turkey, in 2004 and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2009. Her Ph.D. research focused on the design, fabrication and characterization of Ge MOSFETs for future technology nodes.

She was a Research Intern with Translucent Inc. in 2006 and with Intel Component Research in 2008. She is currently with the Center for Integrated Systems and the Department of Electrical Engineering, Stanford University. She is currently working on novel memory and storage devices and nanoscale electronic devices for neuromorphic applications.

Dr. Kuzum was a recipient of a number of awards, including a Texas Instruments Fellowship and an Intel Foundation Fellowship.

**H.-S. Philip Wong** (S'81–M'82–SM'95–F'01) received the B.Sc.(Hons.) degree in electrical engineering from the University of Hong Kong, Pokfulam, Hong Kong, in 1982, the M.S. degree in electrical engineering from the State University of New York at Stony Brook, Stony Brook, in 1983, and the Ph.D. degree in electrical engineering from Lehigh University, Bethlehem, PA, in 1988.

He joined the IBM T. J. Watson Research Center, Yorktown Heights, NY, in 1988. While at IBM, he worked on CCD and CMOS image sensors, double-gate/multigate MOSFET, device simulations for advanced/novel MOSFET, strained silicon, wafer bonding, ultrathin body SOI, extremely short gate FET, germanium MOSFET, carbon nanotube FET, and phase change memory. He held various positions from Research Staff Member to Manager and Senior Manager. While he was a Senior Manager, he had the responsibility of shaping and executing IBM's strategy on nanoscale science and technology, as well as exploratory silicon devices and semiconductor technology. In September 2004, he joined Stanford University as a Professor of electrical engineering, being affiliated with the Center for Integrated Systems and the Department of Electrical Engineering, Stanford University. His research interests are in nanoscale science and technology, semiconductor technology, solid-state devices, and electronic imaging. He is interested in exploring new materials, novel fabrication techniques, and novel device concepts for future nanoelectronics systems. Novel devices often enable new concepts in circuit and system designs. His research also includes explorations into circuits and systems that are device-driven. His current research covers a broad range of topics, including carbon nanotubes, semiconductor nanowires, self-assembly, exploratory logic devices, nanoelectromechanical devices, novel memory devices, and biosensors.

Dr. Wong served on the IEEE Electron Devices Society as elected Administrative Committee member from 2001 to 2006. He served on the IEEE International Electron Devices Meeting (IEDM) committee from 1998 to 2007 and was the Technical Program Chair in 2006 and the General Chair in 2007. He served on the International Solid-State Circuits Conference (ISSCC) program committee from 1998 to 2004 and was the Chair of the Image Sensors, Displays, and MEMS subcommittee from 2003 to 2004. He serves on the Executive Committee of the Symposia of VLSI Technology and Circuits. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON NANOTECHNOLOGY in 2005–2006. He is a Distinguished Lecturer of the IEEE Electron Devices Society (since 1999) and the Solid-State Circuit Society (2005–2007).