Mimicking the Spike-Timing Dependent Plasticity in HfO₂-based Memristors at Multiple Time Scales

M. Maestro-Izquierdo*, M.B. Gonzalez and F. Campabadal

Institut de Microelectrònica de Barcelona, IMB-CNM (CSIC), Campus UAB, 08193 Bellaterra, Spain

**Corresponding author: marcos.maestro@imb-cnm.csic.es*

Abstract

In this work, TiN/Ti/HfO₂/W memristors have been investigated to mimic the spike-time dependent plasticity (STDP) of biological synapses at multiple time scales. For this purpose, a smart software tool has been implemented to control the instrumentation and to perform a dedicated ultra-fast pulsed characterization. Different time scales, from tens of milliseconds to hundreds of nanoseconds, have been explored to emulate the STDP learning rule in electronic synapses. The impact of such times on the synaptic weight potentiation and depression characteristics has also been discussed.

Keywords: Electronic Synapses, HfO₂, Memristor, Resistive Switching, Spike-Timing Dependent Plasticity (STDP), Time Scale

1. Introduction

Artificial Intelligence Systems based on brain-inspired (neuromorphic) computation are an emergent field that aims to develop electronic systems capable of emulating the neuronal learning mechanisms [1], [2].

Biologically, a neural network is formed by approximately 10^{11} neurons interconnected by 10^{15} synapses (Fig. 1) being the estimated energy consumption per synaptic event of ~ 100 fJ [3]. Generally, a neuron transports the information as electrical impulses, known as action potential or spikes. However, sharing this information with the other interconnected neuron is not a straightforward process: here it is where synapses come into play. The generation and transmission of an action potential from the pre-synaptic neuron is weighted by the synaptic strength and applied to the post-synaptic neuron through the release of the neurotransmitters, i.e., chemicals released from a neuron (Fig. 1). When the total input of a neuron (received from all the connected synapses) exceeds a threshold, then the post-synaptic neuron will fire and transmit an output signal. The amount of neurotransmitters released into the synapse is quantified by the synaptic weight that

can be strengthened or weakened, and the ability of a synapse to modify the synaptic weight is known as synaptic plasticity [3].

In neuromorphic systems, neural network elements, neurons and synapses, have been typically implemented by means of CMOS circuits [4]. However, the need for a very large connectivity between neurons (or synapses) makes the use of transistors in reasonably sized circuits difficult.

The emergence of nanoscaled memristor devices with great scalability and low power consumption has unveiled the possibility of replacing CMOS circuits to emulate the behavior of real synapses [5], [6], [7]. Among them, those based on filamentary electrochemical reactions are currently being explored as electronic synapses [5], [6], as they present similarities with biological synapses regarding the mechanisms taking place (Fig. 1). To act as biological synapses, however, memristors must also mimic the synaptic learning rules [8], [9]. One of the most studied synaptic behaviors is the so-called spike-timing dependent plasticity (STDP) [8], [10], which depends on the arrival time of the spikes from pre and post-synaptic neurons. The STDP rule adjusts the connection forces between neurons (synaptic weights) according to the relative time

Fig. 1: (From top to bottom): Schematic representation of a biological connection between two neurons through a synapse. Sketch of a neurotransmitter transfer process in a biological synapse. Structure of the valence change-based memristors employed in this work to mimic the synaptic behavior.

between the output and input potentials of a neuron i.e., the synaptic weight can be strengthened (potentiated) or weakened (depressed).

Several works have investigated the suitability of memristors fabricated with different material combinations [2], [11]-[15] and particularly those based on HfO_2 [12], [13], [15] to act as electronic synapses by analyzing their synaptic potentiation and depression characteristics. Although significant progress has been achieved, the implementation of memristors as artificial synapses is still a challenge even at device level. Moreover, the action potential shape and time should be carefully designed as these parameters could strongly affect the device energy consumption and the STDP response [16], [17].

In this work, the impact of the spike time scale on the STDP behavior of memristive synapses is investigated. For this purpose, HfO_2 -based memristors have been fabricated and electrically characterized. After that, the analog control of memristor conductivity has been proved thus emulating the connectivity between neurons. Then the STDP has been experimentally demonstrated by employing pre- and post-spikes in the range of milliseconds (as biological spikes). Finally, the

impact of different time scales on STDP characteristics (potentiation and depression) is discussed.

2. Samples description

The memristor devices fabricated in this work are $\text{TiN}/\text{Ti}/\text{HfO}_2/\text{W}$ cross point structures with an area of $5 \times 5 \mu\text{m}^2$. They were processed on silicon wafers on a thermally grown 200nm-thick silicon dioxide. Top and bottom electrodes were deposited by magnetron sputtering and patterned by photolithography and dry etching. The top electrode consists of a 200nm-TiN layer on a 10nm-Ti layer acting as an oxygen reservoir, and the bottom electrode is a 200nm-thick W layer. The 10nm-thick HfO_2 insulator was grown by atomic layer deposition at 225°C using TDMAH and H_2O as precursors and N_2 as carrier and purge gas. Finally, a contact area to the bottom electrode was defined by photolithography and dry etching of the HfO_2 film on the W layer.

3. Measurement procedure

A block diagram of the experimental setup used for all the measurements presented in this work is shown in Fig. 2. The measurement setup consists of a laptop, two remote sense/switch units (RSU) that allow switching between DC and AC measurements, and a semiconductor parameter analyzer (SPA), Keysight B1500. The SPA is equipped with two specific modules. The B1511B medium power source measurement unit (MPSMU) module used for quasi-static current-voltage (I-V) measurements, and the B1530 module, a waveform generator/fast

Fig. 2: Block diagram of the measurement setup for ultra-fast pulsed measurements.

measurement unit (WGFMU), that allows measuring ultra-fast pulsed current-voltage characteristics synchronized with the applied waveform with up to 5ns sampling speed. For the measurements, a software tool implemented in C++ was developed to control all the instrumentation via GPIB and to perform sequential and smart measurements.

The flow diagram of the STDP measurement procedure is shown in Fig. 3. First, GPIB communication with the equipment is established. Then, the device is initially driven to an intermediate conductance state (G_{INITIAL}). To do that, the MPSMU module applies, with RSU at DC mode, double sweeps of negative and positive voltages from 0V to -1.5V and 0.9V, respectively. Once the device is set at G_{INITIAL} after the last positive voltage ramp, the RSU are switched to AC mode and two identical and symmetric spike waveforms, corresponding to pre- and post-synaptic spikes, are generated within the WGFMU. The spikes are applied to the TiN and W electrodes, respectively, with a given relative time ($\Delta t = t_{\text{POST}} - t_{\text{PRE}}$) to be tested. To verify the conductance change of the device, its final state is measured applying a -0.1V square pulse. After that, the next Δt values are tested with the same procedure. This sequence is repeated 18 times in which the post-spike is delayed (negative Δt) or ahead (positive Δt)

Fig. 3: Flow diagram of the STDP measurement procedure.

with respect to the pre-spike. Finally, the whole procedure is repeated for 6 different spike widths (in the range from 100ms to 1μs).

4. Results and discussion

The fabricated TiN/Ti/HfO₂/W memristors show bipolar resistive switching characteristics [18]. Typical current-voltage (I-V) RS characteristics are shown in Fig. 4. This behavior is attributed to the formation and rupture of an oxygen vacancy-based conductive filament (CF) in the dielectric layer. Initially, the CF is electroformed applying a positive voltage ramp to the top electrode with a current limitation of 1mA. During such a process, the device conductance is drastically increased. Subsequently, by applying negative and positive voltages, respectively, allow the memristor to change between different conductance states due to morphological and stoichiometric variations of the conductive filament. However, when positive voltages are applied, current must not exceed a critical level to prevent irreversible dielectric breakdown. Therefore, either a current compliance limit or a maximum voltage threshold voltage is required to be considered during the measurements. In this work, the applied voltage to the device was lower than 2V.

In order to evaluate the capability of the fabricated devices to reach different conductance states, the devices were subjected to double sweeps fixing a maximum voltage value in each polarity, $V_{\text{STOP+}}$ and $V_{\text{STOP-}}$. While for $V_{\text{STOP-}}$ only one value (-1.5V) was used, $V_{\text{STOP+}}$ was varied from 0.5V to

Fig. 4: Examples of experimental bipolar RS I-V curves where $V_{\text{STOP+}}$ is varied from 0.5V to 2V with $V_{\text{STOP-}}$ fixed at -1.5V.

2V, allowing devices to reach different current values.

From the experimental I-V curves, the current at -0.1V for each V_{STOP+} condition was extracted. The results are depicted in Fig. 5, showing that the memristor conductivity presents an analog-like behavior with the values spanning two orders of magnitude, i.e., the conductivity of the device can be continuously selected depending on the applied voltage. These results show the capability of the devices to tune the synaptic weight.

After having demonstrated that the memristor conductivity can be electrically tuned, the STDP was experimentally verified in our devices following the measurement procedure described in Fig. 3. Fig. 6(a) shows the used pre-spike shape (blue) and the whole set of the related delayed and ahead post-spikes (red, orange and green). The differences between pre-spike and post-spikes times result in square voltage pulses that cause the device to switch to a final conductance state (G_{FINAL}). When Δt is negative (post-spike preceding the pre-spike), a square pulse with negative voltage amplitude is applied leading to a synaptic weight depression. Fig. 6(b) and (d) show examples of that type of pulse for the largest and shortest tested relative times, respectively. Contrarily, a resultant square pulse with positive voltage amplitude is applied if Δt is positive (pre-spikes preceding the post-spike), leading to synaptic weight potentiation. Fig. 6(c) and (e) show examples of the resultant positive square pulses for the largest

Fig. 5 Current values measured at -0.1V for several V_{STOP+} conditions obtained from the experimental I-V curves. A continuous state selectivity is obtained, spanning two orders of magnitude.

and shortest tested relative times, respectively. Finally, obtaining the initial and final conductance states at the different positive and negative Δt , the STDP is reproduced.

Fig. 7 shows the experimental STDP depicted as the relative conductance variation $\Delta G / G_{INITIAL}$ as a function of the relative times, being $\Delta G = G_{FINAL} - G_{INITIAL}$. Insets in Fig. 7 indicate the resultant square pulses applied at the conditions in Fig. 6(b)-(e). Notice that a square pulse with negative amplitude causes the device conductance to decrease (synaptic depression), whereas, a square pulse with positive amplitude makes the device to increase its conductance (synaptic potentiation). The shorter the $|\Delta t|$, the larger the potentiation and depression. It should be noticed that the observed asymmetry of the STDP, i.e. a larger change under depression than under potentiation, depends on initialization condition, since the range of available states in each condition (potentiation or depression) could be different depending on which the initial state is. With

Fig. 6: (a) Pre- and post- spikes applied to device terminals to reproduce experimentally the STDP learning rule. Several spike widths ranging from 1μs to 100ms were tested. Positive and negative time delays, $\Delta t = t_{POST} - t_{PRE}$, are indicated. Resultant square spikes at (b) the longest negative tested Δt , (c) the longest positive Δt , (d) the shortest negative tested Δt and (e) shortest positive Δt .

Fig. 7: Example of experimental STDP when sequences of pre-spike and post-spike pairs in the range of milliseconds are applied. Insets show the examples of the resultant square pulses in Fig. 6(b)-(e) that give rise to larger or smaller depression or potentiation in the device.

the STDP reproduced in our HfO₂-based memristors, the impact of different time scales on STDP characteristics is next analyzed.

The effect of the spike time scale on the STDP characteristics was explored following the measurement procedure shown in Fig. 3 and varying the spike widths from 100ms to 1μs during the waveform generation step. Therefore, six different time scales of the relative timing were investigated (from 10ms to 100ns). Fig. 8 shows six consecutive STDP measurements at the different time scales performed in the same device (Fig. 8(a) to Fig. 8(f)). The experimental behavior was analyzed according to the biological trend reported [17], [19] and considering the exponential behavior for the synaptic weight update following the Eq. 1 (fitting lines in Fig. 8).

$$\begin{cases} \frac{\Delta G}{G_{\text{INITIAL}}} = +A_+ e^{-\Delta t/\tau_+} \text{ for } \Delta t > 0 \\ \frac{\Delta G}{G_{\text{INITIAL}}} = -A_- e^{+\Delta t/\tau_-} \text{ for } \Delta t < 0 \end{cases} \quad (1)$$

In this equation, time parameters (τ_+ and τ_-) determine the ranges of relative times in which

Fig. 8: Experimental STDP for six different spike widths: (a) 100ms, (b) 10ms, (c) 1ms, (d) 100μs, (e) 10μs and (f) 1μs. Dots indicate the experimental data while lines indicate the fitting of the experimental results according to Eq. 1.

potentiation and depression occur, while scaling parameters (A_+ and A_-) indicate the maximum synaptic modification with respect to the initial state during synaptic potentiation and depression functions, respectively.

Fitting the experimental results in Fig. 6, the characteristic parameters in Eq. (1) can be extracted. Fig. 9 shows the obtained time parameters as a function of the time scale of the relative timing (Δt). Notice that both τ_+ and τ_- increase with the time scale (Fig. 9(a) and (b), respectively), where similar values are obtained for both parameters (Fig. 9(c)). Fig. 10 shows the scale parameters, A_+ , A_- and A_-/A_+ ratio as a function of the time scale. The experimental results indicate that both A_- and A_+ parameters show only slight variations with time and that no clear trend is observed.

Fig. 9: Time fitting parameters of Eq. 1, (a) τ_+ , (b) τ_- and (c) τ_-/τ_+ ratio, as a function of the time scale of Δt .

Fig. 10: Scale fitting parameters of Eq. 1, (a) A_+ , (b) A_- and (c) A_-/A_+ ratio as a function of the time scale of Δt .

5. Conclusions

In this work, the capability of varying the conductivity of HfO_2 -based artificial synapses has been experimentally demonstrated. The spike-timing dependent plasticity (STDP) learning rule of biological synapses has been validated in these devices, indicating their capability to tune their synaptic weight. In addition, the impact of the spike width on the STDP response has been explored. The experimental results have been analyzed according to the biological trend, and considering the exponential behavior for the synaptic weight update function, where the A_+ , A_- , τ_+ and τ_- parameters have been extracted. These parameters have shown linear (τ_+ and τ_-) and constant (A_- and A_+) dependencies on time. The results demonstrate that the STDP can be successfully implemented at different time scales,

indicating the potential of these devices to minimize the energy consumption per synaptic event.

Acknowledgements

This work has been partially supported by the Spanish Ministry of Science, Innovation and Universities and the FEDER program through project TEC2017-84321-C4-1-R. It has made use of the Spanish ICTS Network MICRONANOFABS.

References

- [1] M. A. Zidan, J.P. Strachan and W.D. Lu, Nat. Electron. 1, (2018) 22-29.
- [2] D. Ielmini, Microelectron. Eng. 190 (2018), 44-53
- [3] S. Yu, Neuro-inspired computing using resistive synaptic devices, Springer, 2017.
- [4] G. Indiveri, E. Chicca and R. Douglas, IEEE Trans. Neural Networ. 17 (2006), 211-221.
- [5] D. Kuzum, S. Yu and H-S P. Wong, Nanotechnology, 24 (2013), 382001.
- [6] S. H. Tan, P. Lin, H. Yeon, S. Choi, Y. Park and J. Kim, APL Materials, 6 (2018), 120901_1-120901_16.
- [7] F. Pan, S. Gao, C. Chen, C. Song and F. Zeng, Materials Science and Engineering R, 83 (2014), 1-59.
- [8] H. Z. Shouval, S. S.-H. Wang and G. M. Wittenberg. Front. Comp. Neurosci., 4:19 (2010), 1-13.
- [9] W. Zhang, Y. Hu, T.-Ch. Chang, K.-Ch. Chang, T.-M. Tsai, H.-L. Chen, et al., IEEE Electron Device Lett. 36:8 (2015), 772-774.
- [10] D.E. Feldman, Neuron, 75:4 (2012), 556-571.
- [11] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder and W. Lu, Nano Lett., 10 (2010), 1297-1301.
- [12] E. Covi, S. Brivio, A. Serb, T. Prodromakis, M. Fanciulli and S. Spiga, IEEE Int. Symp. on Circuits and Systems (ISCAS) May 2006, pp. 393-396.
- [13] Y. Matveyev, R. Kirtaev, A. Fetisova, S. Zakharchenko, D. Negrov and A. Zenkevich, Nanoscale Research Letters, 11:147 (2016), 1-6.
- [14] W. Banerjee, Q. Liu, H. Lv, S. Long and M. Liu, Nanoscale 9 (2017), 14442-14450.
- [15] A. Padovani, J. Woo, H. Hwang and L. Larcher, IEEE Electron Dev. Letters, 39:5 (2018), 672-675.
- [16] K. A. Campbell, K. T. Drake and E. H. Barney Smith, Front. Bioeng. Biotechnol., 4:97 (2016), 1-11.
- [17] C. Zamarreño-Ramos, L. A. Camuñas-Mesa, J. A. Pérez-Carrasco, T. Masquelier, T. Serrano-Gotarredona and B. Linares-Barranco, Front. Neurosci., 5:26 (2011), 1-22.

- [18] S. Poblador, M.B. Gonzalez and F. Campabal, Microelectronic Engineering, 187-188 (2018), pp. 148-153.
- [19] S. Song, K. D. Miller and L. F. Abbott, Nat. Neurosci., 3:9 (2000), 919–926.