ARTICLES

Journal of Semiconductors
(2021) 42, 064101
doi: 10.1088/1674-4926/42/6/064101

Oscillation neuron based on a low-variability threshold switching device for high-performance neuromorphic computing

Yujia Li[1, 2], Jianshi Tang[2, 3, †], Bin Gao[2, 3], Xinyi Li[2], Yue Xi[2], Wanrong Zhang[1], He Qian[2, 3], and Huaqiang Wu[2, 3, †]

[1]Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

[2]Institute of Microelectronics, Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China

[3]Beijing Innovation Center for Future Chips (ICFC), Tsinghua University, Beijing 100084, China

Abstract: Low-power and low-variability artificial neuronal devices are highly desired for high-performance neuromorphic computing. In this paper, an oscillation neuron based on a low-variability Ag nanodots (NDs) threshold switching (TS) device with low operation voltage, large on/off ratio and high uniformity is presented. Measurement results indicate that this neuron demonstrates self-oscillation behavior under applied voltages as low as 1 V. The oscillation frequency increases with the applied voltage pulse amplitude and decreases with the load resistance. It can then be used to evaluate the resistive random-access memory (RRAM) synaptic weights accurately when the oscillation neuron is connected to the output of the RRAM crossbar array for neuromorphic computing. Meanwhile, simulation results show that a large RRAM crossbar array (> 128 × 128) can be supported by our oscillation neuron owing to the high on/off ratio (> 10^8) of Ag NDs TS device. Moreover, the high uniformity of the Ag NDs TS device helps improve the distribution of the output frequency and suppress the degradation of neural network recognition accuracy (< 1%). Therefore, the developed oscillation neuron based on the Ag NDs TS device shows great potential for future neuromorphic computing applications.

Key words: threshold switching; Ag nanodots; oscillation neuron; neuromorphic computing

Citation: Y J Li, J S Tang, B Gao, X Y Li, Y Xi, W R Zhang, H Qian, and H Q Wu, Oscillation neuron based on a low-variability threshold switching device for high-performance neuromorphic computing[J]. *J. Semicond.*, 2021, 42(6), 064101. http://doi.org/10.1088/1674-4926/42/6/064101

1. Introduction

A resistive random-access memory (RRAM)-based neural network has been extensively studied as a promising solution to overcome the von Neumann bottleneck faced in conventional artificial intelligence (AI) hardware[1–4]. As inspired by a biological neural network, an artificial neural network consists of synaptic and neuronal devices. In order to improve speed and power efficiency, the RRAM crossbar array, which can significantly accelerate the vector-matrix multiplication, has been developed to implement artificial synapses[5–8]. On the other hand, a neuronal device is needed at the end of each crossbar bit line (BL) to convert the weighted sum current of the analog RRAM synapses into spikes to transmit information to the next layer of neurons. Here integrate-and-fire neurons built with CMOS circuits are typically employed[9]. However, such complex CMOS neurons would occupy a much larger footprint than the BL pitch of the crossbar array, which causes serious column pitch matching problem[10, 11].

Recently, a compact oscillation neuron based on a metal–insulator transition (MIT) threshold switching (TS) device was proposed as a more scalable artificial neuron[12–16]. Compared to the CMOS neuron, an oscillation neuron has the benefits of small size and simple circuit structure, which is appealing for large-scale neuromorphic system applications. However, the on/off ratio of a typical MIT TS device is small (~10^2), which cannot be used for a large RRAM crossbar array (e.g., 12 × 1 array)[15]. Moreover, the high operation voltage of the MIT TS device may disturb the weights of the RRAM synapses and also increase the power consumption[16]. Alternatively, TS devices based on electrochemical metallization (ECM) filaments have attracted considerable attention due to their simple structure, large on/off ratio, and low operation voltage[17–23]. However, the uniformity of typical ECM TS devices is relatively poor, which may affect the accuracy of artificial neural networks. More recently, we have developed a high-uniformity HfO_2-based TS device with patterned Ag nan-

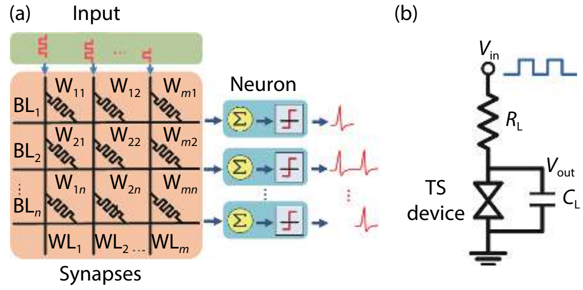


Fig. 1. (Color online) (a) Schematic diagram of a typical artificial neural network. (b) Circuit implementation of the oscillation neuron with a TS device.

Correspondence to: J S Tang, jtang@tsinghua.edu.cn; H Q Wu, wuhq@tsinghua.edu.cn

Received 6 JANUARY 2021; Revised 8 FEBRUARY 2021.

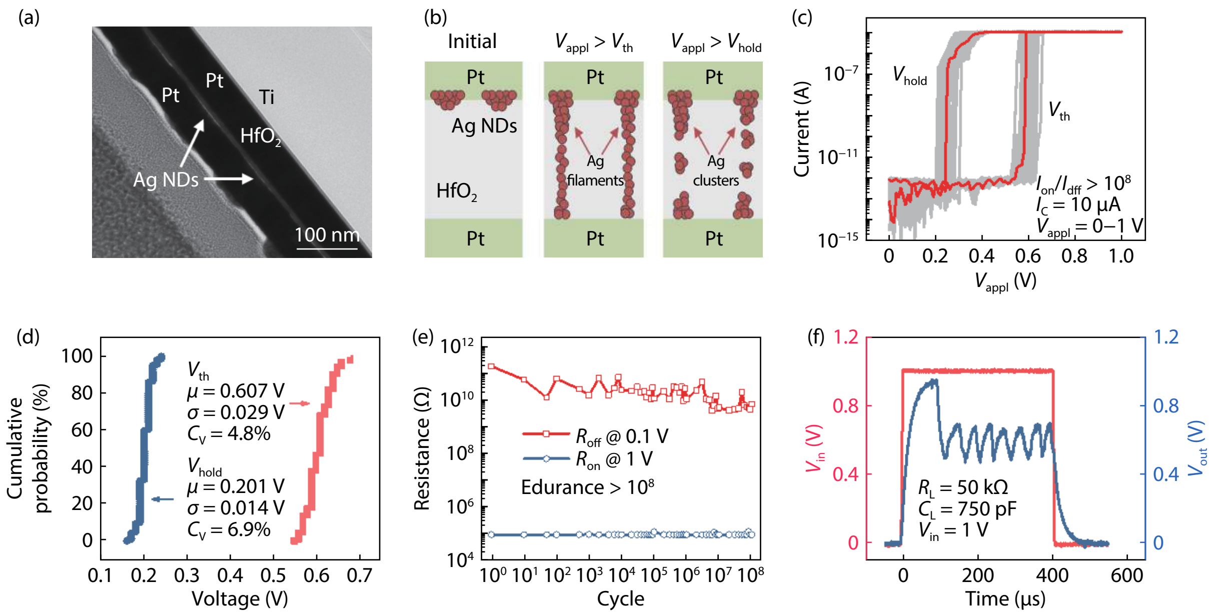


Fig. 2. (Color online) (a) TEM image of the Ag NDs TS device. (b) Schematic illustration of the threshold switching process in the device. (c) Typical current–voltage (I–V) curves for the Ag NDs TS device. (d) Cumulative probability of V_{th} and V_{hold} distributions for the Ag NDs TS device. (e) Endurance test of the Ag NDs TS device with over 10^8 cycles. (f) Measured oscillation waveform of the oscillation neuron.

odots (NDs) as the high-performance selector, which shows low leakage current (< 1 pA), high on/off ratio (> 10^8), and high endurance (> 10^8 cycles)[24].

In this work, we further implement an oscillation neuron using the HfO_2/Ag NDs TS device. This neuron exhibits self-oscillation behavior at low applied voltage (1 V), where the oscillation frequency increases with the applied voltage and decreases with the load resistance. In addition, it can work with a large RRAM crossbar array (> 128 × 128) owing to the high on/off ratio (> 10^8) of Ag NDs TS device. Moreover, in the neural network simulation, a high recognition accuracy (loss < 1%) can be achieved by using this oscillation neuron because of its high uniformity.

2. Results and discussion

Fig. 1(a) illustrates the schematic diagram of a simple artificial neural network. When the input voltages are applied to the crossbar synaptic array, the weighted sum currents are integrated by the neurons at the end of each column (BL) and trigger output spike firing when they reach the neuron thresholds. The circuit implementation of the oscillation neuron based on the TS device is shown in Fig. 1(b). The load resistance (R_L) represents the RRAM synaptic weight connected in series with the TS device. Also, C_L is the load capacitance including parallel capacitance and parasitic capacitance at the neuron node. Initially, the TS device is in the off state (R_{off}). When applying an input voltage pulse (V_{in}), the voltage mainly drops on the TS device since $R_{off} > R_L$, and C_L starts to charge. When V_{in} is larger than the threshold voltage of the TS device V_{th}, it turns to the on state (R_{on}), and then C_L starts to discharge since the voltage drop on the TS device is reduced ($R_{on} < R_L$). Once the voltage drop on the TS device is below the hold voltage of TS device V_{hold}, it switches back to R_{off}, ready for the next firing event. In this way, the TS device switches on and off between R_{on} and R_{off}, and the neuron outputs oscillation signal. If R_L is chosen to satisfy $R_{off} \gg R_L \gg R_{on}$, the ideal oscillation frequency f can be described as[13]:

$$f = 1 \Big/ \left[R_L C_L \times \log\left(\frac{V_{hold} - V_{in}}{V_{th} - V_{in}}\right)\right]. \quad (1)$$

Eq. (1) shows a one-to-one correspondence between f and R_L at a certain C_L and V_{in}. Therefore, the oscillation frequency can be used to represent the weight of the RRAM synapse accurately.

In this study, we demonstrate the oscillation neuron using the Ag NDs TS device based on ECM filaments. The devices were fabricated with a cell size of 10 × 10 μm^2. The bottom electrode was patterned by photolithography and deposited by sputtering of 5 nm Ti and 50 nm Pt. An 8 nm thick HfO_2 dielectric was deposited by atomic layer deposition (ALD) at 250 °C. The Ag NDs with diameters of 50 nm were patterned by e-beam lithography (EBL) and deposited by sputtering. A 40 nm-thick Pt was deposited as the top electrode. The transmission electron microscope (TEM) image of the Ag NDs TS device is shown in Fig. 2(a). Fig. 2(b) exhibits the schematic illustration of the threshold switching process in the device. In the initial state, there is no conductive filament in the dielectric layer, and the device is in the high-resistance state (HRS). When applying a voltage (V_{appl}) larger than V_{th}, the electric field is locally enhanced in the areas with Ag NDs, so the Ag atoms are easier to be ionized as the ion source ($Ag \rightarrow Ag^+ + e^-$) for diffusing toward the bottom electrode, which leads to the formation of metallic filaments and turns the device to a low-resistance state (LRS). In this device, the Ag filaments are thin and unstable due to the small amount of Ag ions. Spontaneous rupture of the filaments occurs immediately when V_{appl} goes below V_{hold}, which turns the device back to HRS. In this process, the Ag atoms form clusters on the trace of filaments. Owing to the highly ordered Ag NDs, Ag filaments tend to form at the same positions and the formed filaments would have similar morphology, in different operation cycles or different devices. Fig. 2(c) shows the typical current–voltage (I–V) curve of the Ag NDs TS device under voltage sweeps between 0 and 1 V. This device exhibits a

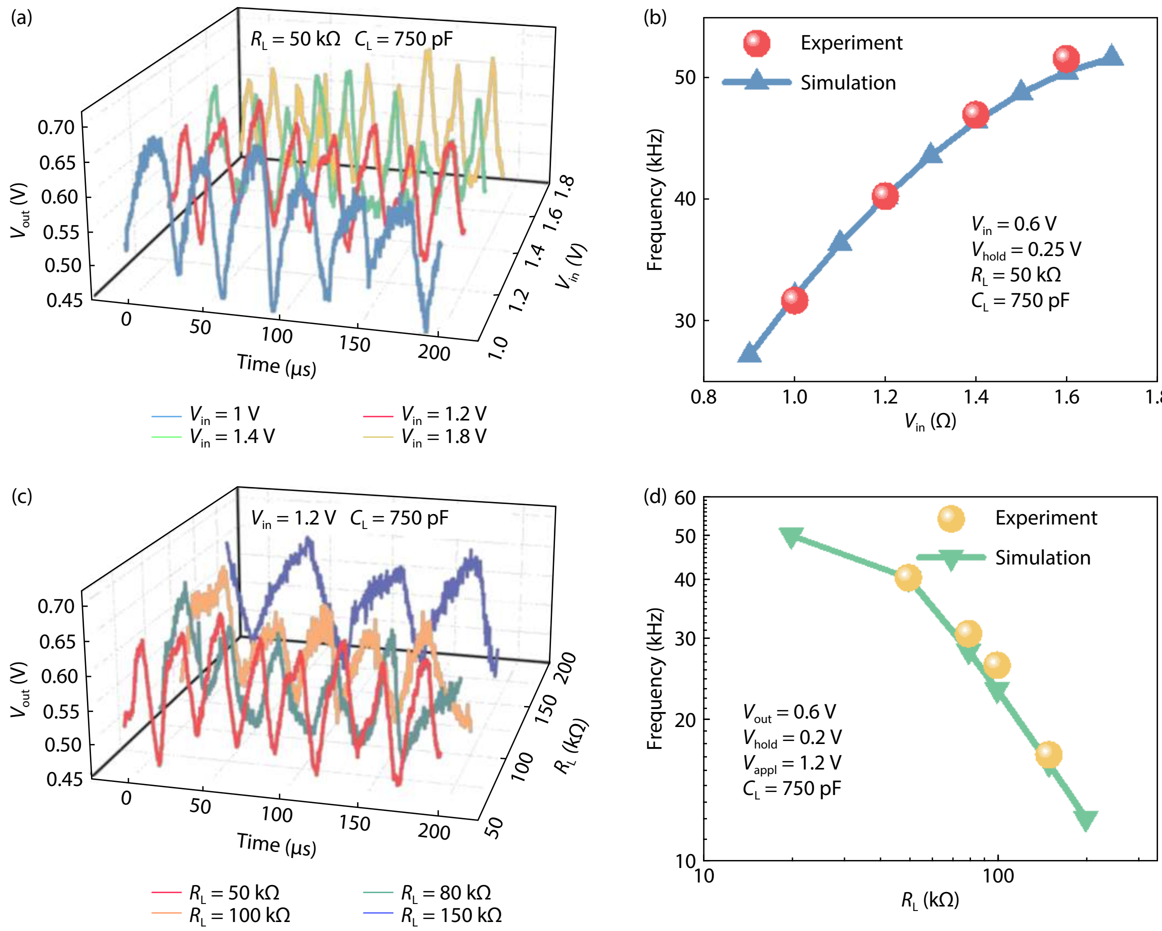


Fig. 3. (Color online) (a) Oscillation waveforms of the oscillation neuron with different V_{in} when R_L = 50 kΩ, C_L = 750 pF. (b) The oscillation frequency as a function of $V_{in.}$ (c) Oscillation waveforms of the oscillation neuron with different R_L when V_{in} = 1.2 V, C_L = 750 pF. (d) The oscillation frequency as a function of $R_{L.}$

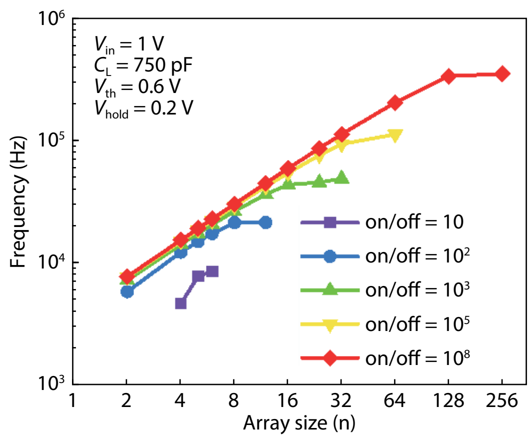


Fig. 4. (Color online) The oscillation frequency as a function of the RRAM crossbar array size under different on/off ratios of the TS device.

low leakage current less than 1 pA and large selectivity over 10^8. The V_{th} and V_{hold} of Ag NDs TS device are 0.6 and 0.2 V, respectively, which are carefully tuned to work with the RRAM synapse. In order to evaluate the uniformity of the Ag NDs TS device, the distributions of V_{th} and V_{hold} are analyzed, and the cumulative probability of V_{th} and V_{hold} is shown in Fig. 2(d). The coefficient of variation is defined as $C_V = \sigma/\mu$ to evaluate the variation, where μ and σ are the mean and standard deviation, respectively. This Ag NDs TS device exhibits excellent uniformity (C_V < 10%) compared to other TS devices based on ECM filaments[24]. The endurance test of the Ag NDs TS device is shown in Fig. 2(e). In this measurement, the device is repeatedly turned on with SET pulses of V_{set} = 1 V and t_{set} = 10 μs, and then relaxed to the off state, which is read with a small pulse of V_{read} = 0.1 V and t_{read} = 10 μs. The on-state current (I_{on}) is limited to 10 μA. It is found that the Ag NDs TS device exhibits a high endurance of over 10^8 cycles. To implement an oscillation neuron, we connect the Ag NDs TS device in parallel with a load capacitance (C_L) and then in series with a load resistor (R_L) following the circuit configuration in Fig. 1(b). Fig. 2(f) shows the measured oscillation waveform of the neuron when R_L = 50 kΩ, C_L = 750 pF and V_{in} = 1 V. The test result shows that this oscillation neuron can output a continuous oscillation signal, when V_{in}, R_L and C_L are fixed. In addition, the Ag NDs TS oscillation neuron shows a certain time delay before its stable oscillation, which leads to a higher V_{out} in the first peak. It is owing to the turn-on delay time of the TS device[24]. The higher first peak may have an adverse effect on the synaptic weight precision by distorting the oscillation waveform, which can be minimized by further improving the TS device switching speed.

More systematic studies on the oscillation neuron characteristics are performed as shown in Fig. 3. Fig. 3(a) shows the oscillation waveforms at different input pulse voltages when R_L = 50 kΩ and C_L = 750 pF. The oscillation frequencies are 31.6, 40.2, 46.9 and 51.4 kHz for V_{in} = 1, 1.2, 1.4, and 1.6 V, re-

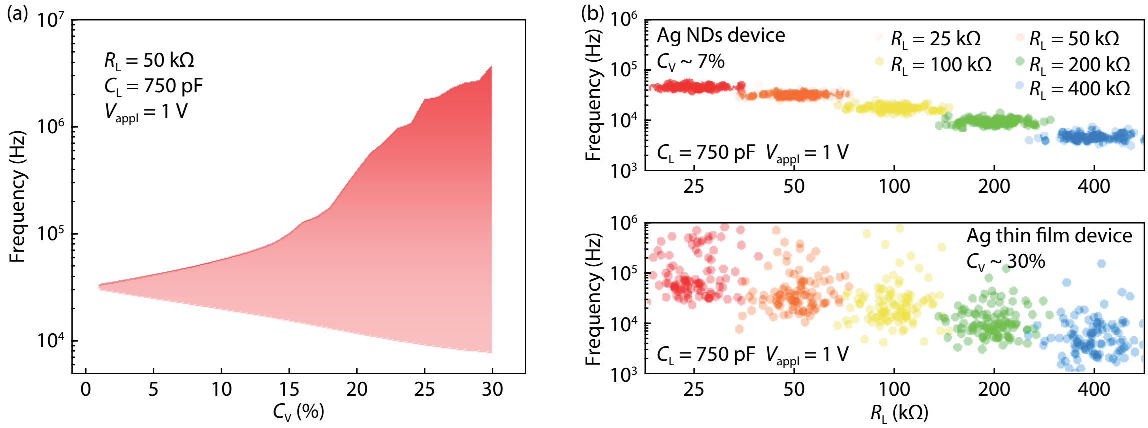


Fig. 5. (Color online) (a) The oscillation frequency distribution under different C_V. (b) The oscillation frequency distribution of different R_L when C_V = 7% (top panel) and C_V = 30% (bottom panel).

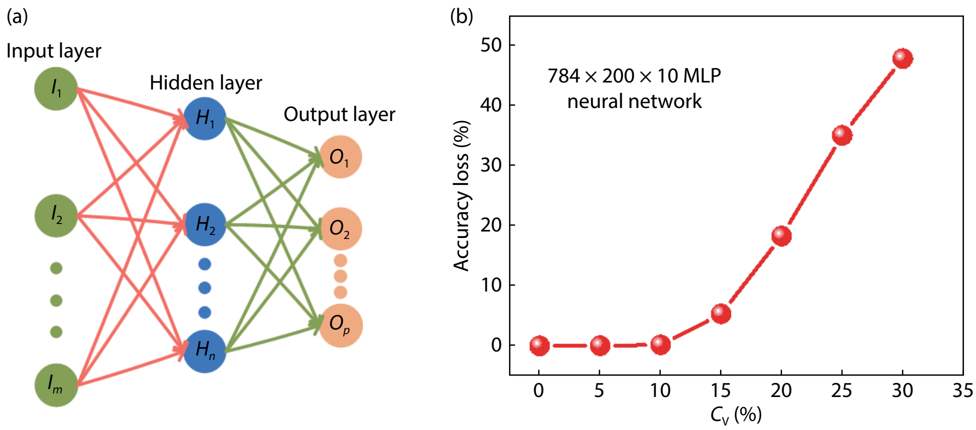


Fig. 6. (Color online) (a) The structure of MLP neural network. (b) Simulation results of the MNIST recognition accuracy loss as a function of the variability of the TS device.

spectively. It is found that the oscillation frequency increases as a function of the pulse amplitude of V_{in} as described in Eq. (1), and shows good consistency with the simulation results, as shown in Fig. 3(b). Similarly, Fig. 3(c) shows the oscillation waveforms of the Ag NDs TS devices with different R_L values (i.e., different synaptic weights) when V_{in} = 1.2 V and C_L = 750 pF. The oscillation frequencies are 40.2, 30.5, 26.1 and 16.8 kHz for R_L = 50, 80, 100 and 150 kΩ, respectively. The oscillation frequency decreases as R_L increases, as shown in Fig. 3(d). The test results indicate that this Ag NDs TS oscillation neuron exhibits self-oscillation behavior, and the output oscillation frequency can be used to sense the weight of the RRAM synapse accurately when the neuron is connected to the crossbar array.

Based on experimentally derived device data, a model of the RRAM crossbar array is developed to evaluate synaptic weights of different array sizes. Then SIPCE simulation is used to investigate the relationship between the oscillation frequency and the RRAM crossbar array size. In order to simplify the simulation process, the resistance RRAM device is fixed at an intermediate resistance state (R_L = 500 kΩ), and V_{in} is applied to all the word lines (WLs) in parallel. The results are shown in Fig. 4. If the on/off ratio of the TS device is less than 100, only a limited range of the weighted sum could meet the criterion for oscillation ($R_{off} > R_L > R_{on}$). It means that the oscillation neuron can only be used for small crossbar arrays (< 16 × 16). With the increase of on/off ratio, the oscillation neuron can work with a wider range of load resistance. Therefore, the weighted sum in a larger RRAM crossbar array can be successfully identified by distinguishable oscillation frequency. As in the case of our present Ag NDs TS device with on/off ratio > 10^8, it can work with a large RRAM crossbar array of size > 128 × 128.

In order to investigate the impact of TS device variation on the oscillation neuron, the oscillation frequency distribution with different C_V under the same experimental conditions is shown in Fig. 5(a). The uniformity of the TS device deteriorates with the increase of C_V, which leads to a large variation of the oscillation frequency under the same load condition. Fig. 5(b) shows the oscillation frequency distribution of different R_L with different TS devices. Compared with the Ag thin film device (C_V ~ 30%)[23], the output frequency distribution of the Ag NDs TS neuron (C_V ~ 7%) is more concentrated. Therefore, the oscillation neuron based on the Ag NDs TS device can achieve higher accuracy for neuromorphic computing.

In order to further analyze the impact of the TS device's uniformity on the artificial neural network, a multi-layer perceptron (MLP) of 784 × 200 × 10 is simulated to classify the handwritten digits in the Modified National Institute of Standards and Technology (MNIST) dataset, as shown in Fig. 6(a)[25]. The simulation results displayed in Fig. 6(b) indicate that the increase in C_V leads to a dramatic degradation of the recognition accuracy, especially when C_V > 15%. Therefore, the oscillation neuron based on the high uniformity Ag NDs TS device developed in this work is beneficial to reduce the network accuracy loss.

3. Conclusion

In conclusion, we have demonstrated a reliable oscillation neuron using the low-variability Ag NDs TS device for high-performance neuromorphic computing. This neuron exhibits self-oscillation behavior at low applied voltages down to 1 V. A systematic study on the oscillation characteristics reveals that the oscillation frequency increases with the applied voltage and also the synaptic conductance connected as the load resistor. The high uniformity and large on/off ratio of the Ag NDs TS device enable the oscillation neuron to reduce the neural network accuracy loss (< 1%) and make it applicable to a large-scale RRAM crossbar array (> 128 × 128). The developed oscillation neuron hence has great potential for future neuromorphic system applications.

Acknowledgements

This work was supported in part by China Key Research and Development Program (2016YFA0201800), and the National Natural Science Foundation of China (91964104, 61974081).

References

[1] Indiveri G, Linares-Barranco B, Legenstein R, et al. Integration of nanoscale memristor synapses in neuromorphic computing architectures. Nanotechnology, 2013, 24, 384010

[2] Ambrogio S, Balatti S, Milo V, et al. Neuromorphic learning and recognition with one-transistor-one-resistor synapses and bistable metal oxide RRAM. IEEE Trans Electron Devices, 2016, 63, 1508

[3] Burr G W, Shelby R M, Sebastian A, et al. Neuromorphic computing using non-volatile memory. Adv Phys X, 2017, 2, 89

[4] Zidan M A, Strachan J P, Lu W D. The future of electronics based on memristive systems. Nat Electron, 2018, 1, 22

[5] Merrikh B F, Prezioso M, Chakrabarti B, et al. Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits. Nat Commun, 2017, 9, 2331

[6] Yao P, Wu H Q, Gao B, et al. Fully hardware-implemented memristor convolutional neural network. Nature, 2020, 577, 641

[7] Choi S, Yang J, Wang G. Emerging memristive artificial synapses and neurons for energy-efficient neuromorphic computing. Adv Mater, 2020, 32, 2004659

[8] Zhu J D, Zhang T, Yang Y C, et al. A comprehensive review on emerging artificial neuromorphic devices. Appl Phys Rev, 2020, 7, 011312

[9] Kadetotad D, Xu Z H, Mohanty A, et al. Parallel architecture with resistive crosspoint array for dictionary learning acceleration. IEEE J Emerg Sel Top Circuits Syst, 2015, 5, 194

[10] Hua Q L, Wu H Q, Gao B, et al. Low-voltage oscillatory neurons for memristor-based neuromorphic systems. Glob Challenges, 2019, 3, 1900015

[11] Dang B J, Liu K Q, Zhu J D, et al. Stochastic neuron based on IGZO Schottky diodes for neuromorphic computing. APL Mater, 2019, 7, 071114

[12] Li S, Liu X J, Nandi S K, et al. High-endurance MHz electrical self-oscillation in Ti/NbO_x bilayer structures. Appl Phys Lett, 2015, 106, 212902

[13] Gao L G, Chen P Y, Yu S M. NbO_x based oscillation neuron for neuromorphic computing. Appl Phys Lett, 2017, 111, 103503

[14] Duan Q X, Jing Z K, Yang K, et al. Oscillation neuron based on threshold switching characteristics of niobium oxide films. 2019 IEEE International Workshop on Future Computing, 2019, 1

[15] Woo J, Wang P N, Yu S M. Integrated crossbar array with resistive synapses and oscillation neurons. IEEE Electron Device Lett, 2019, 40, 1313

[16] Wang P N, Khan A I, Yu S M. Cryogenic behavior of NbO_2 based threshold switching devices as oscillation neurons. Appl Phys Lett, 2020, 116, 162108

[17] Luo Q, Xu X, Lv H, et al. Cu BEOL compatible selector with high selectivity (> 10^7), extremely low off-current (pA) and high endurance (> 10^{10}). 2015 IEEE International Electron Devices Meeting (IEDM), 2015, 10.4.1

[18] Yoo J, Woo J, Song J, et al. Threshold switching behavior of Ag-Si based selector device and hydrogen doping effect on its characteristics. AIP Adv, 2015, 5, 127221

[19] Du G, Wang C, Li H X, et al. Bidirectional threshold switching characteristics in Ag/ZrO_2/Pt electrochemical metallization cells. AIP Adv, 2016, 6, 085316

[20] Wang Z R, Rao M Y, Midya R, et al. Threshold switching: Threshold switching of Ag or Cu in dielectrics: Materials, mechanism, and applications. Adv Funct Mater, 2018, 28, 1870036

[21] Yoo J, Park J, Song J, et al. Field-induced nucleation in threshold switching characteristics of electrochemical metallization devices. Appl Phys Lett, 2017, 111, 063109

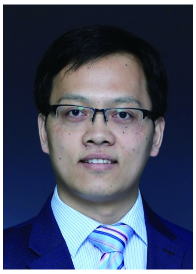
[22] Wang W, Wang M, Ambrosi E, et al. Surface diffusion-limited lifetime of silver and copper nanofilaments in resistive switching devices. Nat Commun, 2019, 10, 81

[23] Hua Q L, Wu H Q, Gao B, et al. Threshold switching selectors: A threshold switching selector based on highly ordered Ag nanodots for X-point memory applications. Adv Sci, 2019, 6, 1970058

[24] Li Y J, Tang J S, Gao B, et al. High-uniformity threshold switching HfO_2 -based selectors with patterned Ag nanodots. Adv Sci, 2020, 7, 2002251

[25] Xi Y, Gao B, Tang J S, et al. In-memory learning with analog resistive switching memory: A review and perspective. Proc IEEE, 2021, 109, 14

Yujia Li is currently a joint PhD. student of Faculty of Information Technology, Beijing University of Technology and Institute of Microelectronics, Tsinghua University from 2016. Her current research interests include design and optimization of resistive switching memory and selector devices as well as their applications in neuromorphic computing.



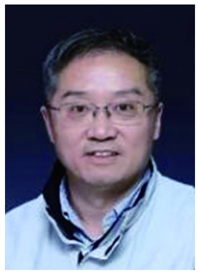
Jianshi Tang (Senior Member, IEEE) received the B.S. degree in microelectronics and nanoelectronics from Tsinghua University, Beijing, China, in 2008, and the Ph.D. degree in electrical engineering from the University of California at Los Angeles, Los Angeles, CA, USA, in 2014. From 2015 to 2019, he worked at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA. He is currently an Assistant Professor with the Institute of Microelectronics, Tsinghua University. He has published over 100 journal articles and conference proceedings, and filed over 60 patents, 20 of which have been granted. His current research interests include emerging memory and neuromorphic computing and carbon nanotube electronics. Prof. Tang has received many awards, including the National Young Thousand Talents Plan, the NT18 Best Young Scientist Award, the IEEE Best Lightning Talk, and the IBM Invention Achievement Awards.

Bin Gao (Senior Member, IEEE) received the B.S. degree in physics and the Ph.D. degree in microelectronics from Peking University, Beijing, China, in 2008 and 2013, respectively. He is currently an Associate Professor of microelectronics with Tsinghua University, Beijing. His current research interests include emerging memory devices and neuromorphic computing.

Xinyi Li received the Ph.D. degree in Microelectronics and Solid State Electronics from TianJin University, TianJin, China, in 2010. She is currently working with the Institute of Microelectronics, Tsinghua University. Her current research interests include neuromorphic devices and their application in neuromorphic computing.

Yue Xi (Student Member, IEEE) received the bachelor's degree in microelectronics from Xi'an Jiaotong University, Xi'an, China, in 2017. He is currently pursuing the Ph.D. degree in microelectronics with Tsinghua University, Beijing, China. His current research interests include ana-log resistive switching memory devices and memristor-based neuromorphic computing.

Wanrong Zhang received the B.S. and M.S. degrees in microelectronics and solid-state electronics from Lanzhou University, Lanzhou, China, in 1987 and 1990, respectively, and the Ph.D. degree in microelectronics and solid-state electronics from Xi'an Jiaotong University, Xi'an, China, in 1996. He is currently a Professor and a Ph.D. Supervisor with the Faculty of Information Technology, Beijing University of Technology, Beijing, China. His current research interests include RF/microwave/millimeter-wave devices and circuits, mixed-signal circuits, cryogenic electronics, device-to-circuit interactions, noise and linearity, reliability physics, device-level simulation, and compact circuit modeling.



He Qian (Member, IEEE) received the Ph.D. degree in microelectronics from Xi'an Jiao-tong University, Xi'an, China, in 1990. From 1990 to 2006, he worked with the Institute of Microelectronics, Chinese Academy of Sciences (IMECAS), Beijing, China, where he became a Professor in 1996 and the Director in 2001. From 2006 to 2008, he worked for the Samsung Semiconductor China Research and Development Center (SSCR), Nanjing, China, as the Director. In 2009, he joined the Institute of Micro-electronics, Tsinghua University (IMTU), Beijing, as a Professor. His current research interests include resistive random access memory (RRAM), 3-D NAND, and neuromorphic computing based on RRAM array.

Huaqiang Wu (Senior Member, IEEE) received the double B.S. degrees in material science and engineering and enterprise management from Tsinghua University, Beijing, China, in 2000, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 2005. From 2006 to 2008, he was a Senior Engineer with Spansion LLC, Sunnyvale, CA, USA. He joined the Institute of Microelectronics, Tsinghua University, in 2009, where he is currently a Professor and the Director of the Institute of Microelectronics. He also serves as the Deputy Director of the Beijing Innovation Center for Future Chip (ICFC). He has published more than 100 technical articles and owns more than 60 patents. His research interests include emerging memories and neuromorphic computing.