Simulation of the Neuromorphic Network Operation Taking into Account Stochastic Effects

Alexander Yu. Morozov[a,b], Karine K. Abgaryan[a,b] and Dmitry L. Reviznikov[a,b]

[a] *Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS), st.Vavilova, 44, bld. 2, Moscow, 119333, Russia*
[b] *Moscow Aviation Institute (National Research University) (MAI), Volokolamskoe Highway, 4, Moscow, 125993, Russia*

Abstract

This paper simulates an analogue self-learning pulse neural network based on memristive elements, taking into account their stochastic properties. A variable-resistor model of a thin-film memristor based on an exponential model for dopant drift is used as a memristor model. Stochastic properties are accounted for by a term in the memristor equation of state, which presents an additive (Gaussian) noise. Memristor switch from low-resistance to high-resistance state in this case occurs differently from cycle to cycle, corresponding to the experimental data. The mathematical model previously developed by the authors is used; it describes the analogue implementation of a pulse self-learning neural network with memristive elements as synaptic weights and a learning mechanism based on the STDP method. The operation of a network consisting of five neurons with 320 synapses for recognition of various black and white images is simulated. As a result of the simulation, the network was successfully learnt to recognize the given patterns

Keywords 1

Memristor, additive noise, neuromorphic network, impulse neural network, STDP, recognition, titanium oxide, TiO2

CEUR-WS.org/Vol-2930/paper10.pdf

1. Introduction

Artificial neural networks play a great role in modern life [1]. With their development, it became possible to study actual and practically significant tasks that often cannot be solved by classical approaches. The recognition task belongs to such tasks. The scope of recognition applications is very extensive: text recognition (including handwriting), machine vision, speech or fingerprints recognition, etc. Neural networks are actively used in such areas as: economics, medicine and healthcare, avionics, the Internet, robotics, security, etc.

One of the factors restraining the neural network development is the high computational complexity of the corresponding neural network algorithms: a network training time can be measured in weeks and months. Currently, to speed up their work, research is being conducted on creating special processors [2] based on the principles of the human brain activity. These processors are often a hardware implementation of pulse neural networks.

In conjunction with the development of specialized computing devices, the use of other computational principles seems promising, namely analogue computations instead of digital ones, since they are performed in orders of magnitude faster. Relatively new electrical elements - memristors [3]-[5] - are actively used in the field of analogue computing. A memristor is a resistor, which conductivity depends on the total electric charge flowing through it. In the absence of current,

VI International Conference Information Technologies and High-Performance Computing (ITHPC-2021), September14–16, 2021, Khabarovsk, Russia
EMAIL: morozov@infway.ru (A. 1); kristal83@mail.ru (A. 2);reviznikov@mai.ru (A. 3)
ORCID: 0000-0003-0364-8665 (A. 1); 0000-0002-0059-0712 (A. 2); 0000-0003-0998-7975 (A. 3)
© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

it keeps its conductivity, which allows it to be used as an elementary memory cell, and it is possible to dynamically change its resistance in the presence of current. This is a certain similarity of the memristors' properties with the properties of biological synapses, allowing them to be used to create analogue neural networks [6].

Basically, the memristive effect occurs in various oxides due to the movement of ions (oxygen vacancies) and the formation/destruction of conducting filaments. The ion movement is random and, as a consequence, memristors have certain stochastic properties. A detailed experimental and theoretical study of this effect was carried out in [7]-[9].

Mathematical models of memristors are traditionally formulated as dynamic systems with respect to the parameter of memristor state, which characterizes the level of element's conductivity [10]-[12]. To account for random effects, a stochastic differential equation with additive noise for the state variable can be used instead of the deterministic one [13].

In this work we simulate a two-layer full-connected network with one layer of memristor elements (synapses) [14]; the network consists of five neurons with 320 synapses. This number of neurons is due to the number of images recognized (5 pieces); the number of synapses is determined by the total number of pixels: 5x8x8 = 320. There is used the network architecture, in which each memristor corresponds to a transistor (1T1R crossbar architecture). Due to this architecture, it is possible to train the network at the hardware level using the STDP (Spike Timing Dependent Plasticity) method [15]-[18]. According to this method, the change of neuron synaptic weights depends on the time difference between input and output pulses.

The main purpose of this scientific work is to simulate the functioning of the self-learning pulse neural network with the hardware implementation based on memristive elements, taking into account their stochastic properties in learning the recognition of five images.

In the second section, a model of a variable-resistor thin-film memristor, based on an exponential model of dopant drift, is presented. The third section provides a mathematical model of the neural network. Next, the neural network operation is simulated and, finally, the main results of the work are formulated.

2. Mathematical model of a memristor

An approach based on representation of a memristor as a dynamic system with a generalized state variable is widely used for simulating the memristor's operation [10-12]. Variation of the state variable determines the dynamics of the element's switching between different modes. Stochastic effects can be accounted for by introducing a stochastic additive term into the dynamic system in the form of additive white (Gaussian) noise [13]. The equation specifying the change in the state variable of the memristor can be written in the following generalized form:

$$dx = F(x)dt + \eta dW,$$

Where F - rate of change of the state variable; η - coefficient characterizing noise intensity; W - Wiener process.

Concretization of the functional dependence $F(x)$ and relation of the state variable with physical parameters gives a memristor model. In the present paper, a variable-resistor thin-film memristor model based on the exponential model for a doping impurity drift is used [19]:

$$dx = \eta dW + \begin{cases} \mu_v \dfrac{V_p}{D^2} \exp\left(\dfrac{R_{on}}{V_p} I\right) dt, V \ge V_p, \\ \mu_v \dfrac{V_n}{D^2} \exp\left(\dfrac{R_{on}}{V_n} I\right) dt, V \le V_n, \\ \mu_v \dfrac{R_{on}}{D^2} I dt, V_n < V < V_p, \end{cases}$$

$$R = R_{on} x + R_{off}(1 - x),$$

$$I = \frac{V}{R},$$

where

$x \in [0,1]$ — state variable;

R_{on}, R_{off} — minimum and maximum memristor resistance;

I, V, R — actual memristor's current, voltage, and resistance values;

V_p, V_n — values of voltages at which state switching occurs;

μ_v —doping mobility coefficient;

D —semiconductor film thickness;

η — coefficient characterizing the noise intensity;

W — Wiener process.

To obtain an approximate realization of the stochastic process $x(t)$ we use the Runge-Kutt method of order 1.5. We determine an approximate solution using the grid $t_0 < t_1 < t_2 < ... < t_{N-1} < t_N$ under the initial condition $x(t_0) = x_0$

$$x_k^1 = x_k + \frac{2}{3}\left[F(x_k)\Delta t_{k+1} + \eta \Delta W_{k+1}\right],$$

$$x_k^2 = x_k + \left[F(x_k^1) - F(x_k)\right]\Delta t_{k+1},$$

$$x_{k+1} = x_k + \frac{1}{4}\left[3F\left(x_k^1\right) + F\left(x_k^2\right)\right]\Delta t_{k+1} + \eta \Delta W_{k+1},$$

where $\Delta t_{k+1} = t_{k+1} - t_k$, $\Delta W_{k+1} = W\left(t_{k+1}\right) - W\left(t_k\right)$. Let $N(0,1)$ is a normally distributed random variable with zero mathematical expectation and unit variance. Then the random value ΔW_{k+1} is calculated as

$$\Delta W_{k+1} = z_{k+1}\sqrt{\Delta t_{k+1}},$$

where z_{k+1} is chosen from $N(0,1)$.

Note that we can use other methods of integrating stochastic ODEs, such as the Milstein method of first order, which for this problem would be equivalent to the Euler-Maruyama method. In this case, the solution on the $(k+1)$-th time layer will be defined as

$$x_{k+1} = x_k + F\left(x_k\right)\Delta t_{k+1} + \eta \Delta W_{k+1}.$$

In this work, the memristor's operation was simulated the following parameter values: $R_{on} = 205$ Ohm, $R_{off} = 2128$ Ohm, $\mu_v = 6 \times 10^{-10}$, $V_p = 0.65$ V, $V_n = -0.87$ V, $D = 621$ nm, $x(0) = 0$, $V(t)$ - see at Fig.1a. Such choice of parameters and voltage form is due to obtaining of memristor characteristics similar to the experimental characteristics for titanium oxide, given in [20]. The current vs. time dependence is shown in Fig. 1b, Fig. 1c shows resistance vs. time dependence, and Fig. 1d shows state change vs. time dependence. Fig. 2 shows the experimental volt-ampere characteristic and the model characteristic.

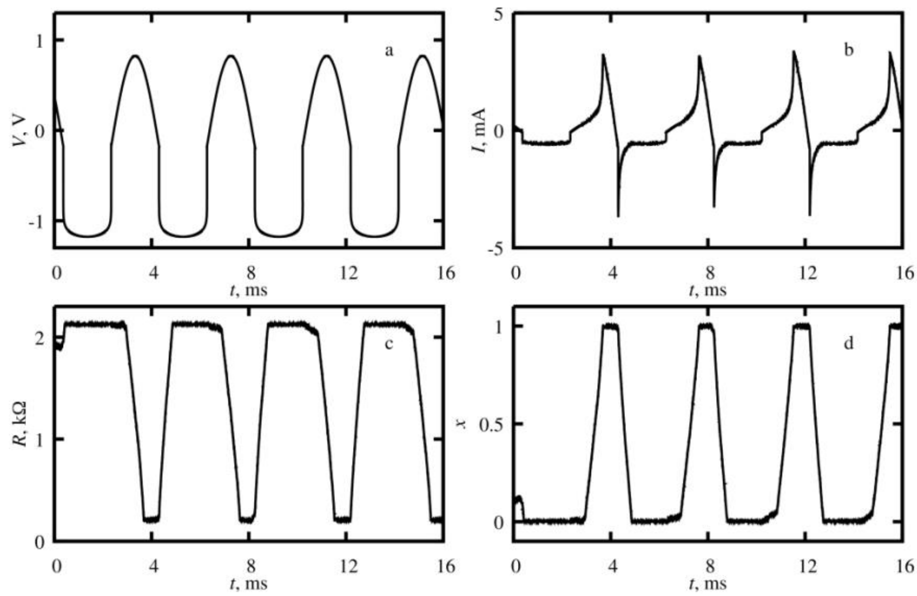

Figure 1: Voltage (a), current (b), resistance (c) and state (d) vs. time dependences

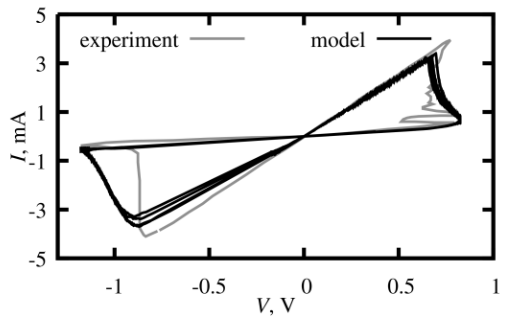

Figure 2: Comparison of the model's volt-ampere characteristic with experimental data for titanium oxide

The presence of noise in the memristor model causes all parameters to take on the stochastic properties. On the diagram with the volt-ampere characteristic, we can distinguish trajectories that correspond to different switching cycles of the memristor. Here we observe the good correspondence in the right part of the diagrams and satisfactory correspondence in their left parts.

3. Mathematical model of a neuromorphic network

We consider a two-layer fully connected self-learning analogue network with one layer of memristor elements (synapses); it consists of 64 inputs and 5 neurons (Fig. 3).

According to the STDP method, the learning mechanism is implemented through feedback (V_{te}). At the moment of neuron activation, two pulses of opposite sign arrive via the feedback channel with delays. If there is activity at the synapse and a positive feedback pulse arrives, then the resistance value of the corresponding memristor decreases, and if a negative feedback pulse arrives, the memristor resistance increases.

The neuron model is a parallel *RC* circuit. As soon as the value of the potential across the capacitor exceeds a certain threshold, its potential is reset, and signals V_{te} and V_{out} are generated. In addition, at the moment of neuron activation, the potential of other neurons decreases in proportion to the coefficient α.

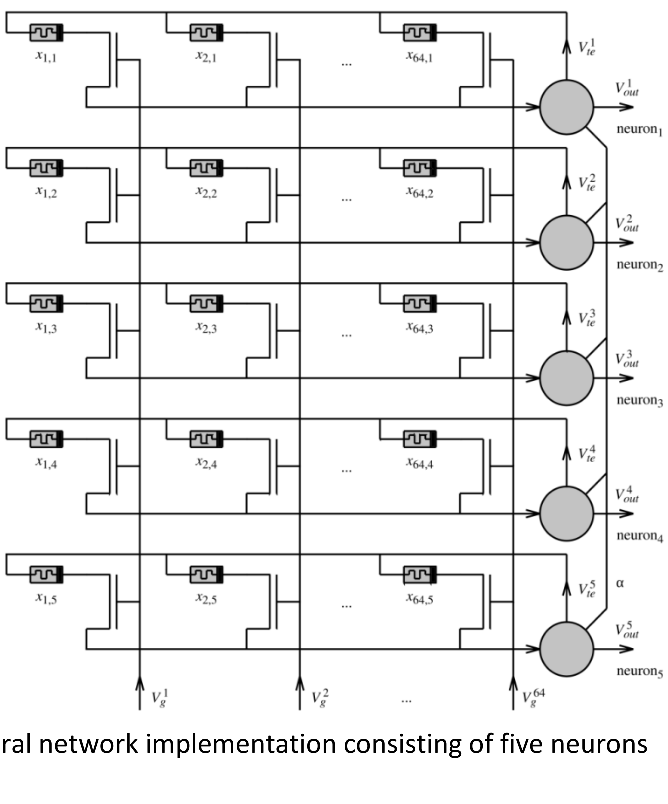

Figure 3: Schematic neural network implementation consisting of five neurons

The mathematical model is set in accordance with [14]. The main difference of the model presented below is the presence of a term ηdW on the right-hand side of the state equations for memristors, which correspond to synapses. As a result, the entire model of the neuromorphic network becomes stochastic

$$dx_{i,j} = \begin{cases} F_X\left(x_{i,j}, V_{te}^j - V_{int}^j\right)dt + \eta dW_{i,j}, V_g^i(t) > 0, \\ 0, V_g^i(t) = 0, \end{cases}$$

$$F_X\left(x_{i,j}, V\right) = \begin{cases} \mu_v \dfrac{V_p}{D^2}\exp\left(\dfrac{R_{on}}{V_p}\dfrac{V}{R_{i,j}}\right)dt, V \geq V_p, \\ \mu_v \dfrac{V_n}{D^2}\exp\left(\dfrac{R_{on}}{V_n}\dfrac{V}{R_{i,j}}\right)dt, V \leq V_n, \\ \mu_v \dfrac{R_{on}}{D^2}\dfrac{V}{R_{i,j}}dt, V_n < V < V_p, \end{cases}$$

$$R_{i,j} = R_{on}x_{i,j} + R_{off}(1 - x_{i,j}),$$

$$\frac{dV_{int}^j}{dt} = \frac{1}{C_{int}}\left[\sum_{\substack{i=1 \\ V_g^i(t)>0}}^{n} \frac{\hat{V}_{te}^j - V_{int}^j}{R_{i,j}} - \frac{V_{int}^j}{R_{int}}\right] - \max_{i=1,m}\left[\theta\left(V_{int}^i - V_{th}\right)\hat{\alpha}_{i,j}\right]\delta\left(\prod_{i=1}^{m}\left(V_{int}^i - V_{th}\right)\right)V_{int}^j,$$

$$\frac{d\tau_j}{dt} = 1 - \delta(V_{int}^j - V_{th})\tau_j,$$

$$V_{out}^j = \begin{cases} V_{out}^+, \tau_j \leq \tau_{out}, \\ 0, \tau_{out} < \tau_j, \end{cases}$$

$$V_{te}^j = \begin{cases} V_{te}^+, \tau_j \leq \tau_s, \\ V_{te}^-, \dfrac{\tau_r}{2} < \tau_j \leq \dfrac{\tau_r}{2} + \tau_s, \\ V_{te}^0, \tau_r < \tau_j, \\ 0, \tau_s < \tau_j \leq \dfrac{\tau_r}{2} \vee \dfrac{\tau_r}{2} + \tau_s < \tau_j \leq \tau_r, \end{cases}$$

$$\hat{\alpha}_{i,j} = 1 - \alpha\left(1 - \delta_{ij}\right), \hat{V}_{te}^j = \max(0, \min(V_{te}^j, V_{te}^0)),$$

$$x_{i,j}(0) = \text{random}\left[0,1\right], V_{int}^j(0) = 0, \tau_j(0) > \max(\tau_r, \tau_{out}), i = 1...64, j = 1...5,$$

where V_g^i is the actual voltage value at the i-th input of the neural network; V_{te}^j - the actual voltage value in the feedback of the j-th neuron; V_{out}^j - the actual voltage value at the output of the j-th neuron; τ_j - time elapsed since the last activation of the j-th neuron; V_{int}^j - voltage across the capacitor of the j-th neuron; R_{int}, C_{int} - the value of resistance and capacitance of neurons; V_{te}^+, V_{te}^-, V_{te}^0 - the amplitude values of the feedback pulses and the default voltage value; V_{out}^+ - the output pulse amplitude; V_{th} - the neuron activation voltage level; $R_{i,j}$ - the memristor's resistance value of the i-th synapse of the j-th neuron; $x_{i,j}$ - the memristor's state of the i-th synapse of the j-th neuron; τ_r - the feedback signal duration after neuron activation; τ_s - the duration of one pulse in the feedback signal, $2\tau_s < \tau_r$; τ_{out} - the duration of one pulse at the network output; α - suppression coefficient; δ_{ij} - the Kronecker symbol; $\delta(x)$ - delta function; $\theta(x)$- Heaviside function; η - coefficient characterizing the noise intensity; $W_{i,j}$- Wiener process corresponding to the i-th memristor of the j-th neuron.

4. Simulation of the neuromorphic network operation. Results

The recognition problem for five patterns is considered (Fig. 4) [21]. The process of training the network is as follows: for each epoch (equal to $\tau_r/2$ seconds), the input signals vector $V_g(t)$ corresponds either to one of the recognized patterns or is set randomly (the elements of the vector have a discrete distribution: V_g^i = 0 V with a probability of 0.73 and V_g^i = 2 V with a probability of 0.27). Over time, the network adapts to pattern recognition. The Distribution of patterns among neurons occurs in the course of training.

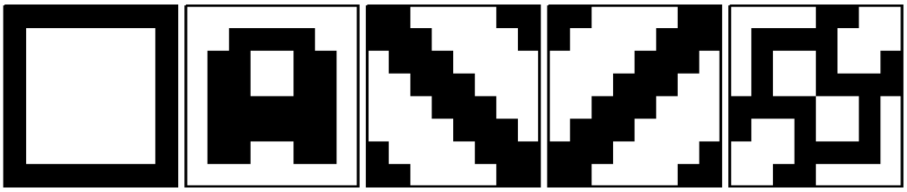

Figure 4: Recognized images

The parameters of the mathematical model of the neural network are adjusted depending on the memristor model. For the used model, which corresponds to a memristor based on titanium oxide (TiO_2), we have the following parameter values: $R_{int} = 200$ Ohms, $C_{int} = 45$ μF, $V_{te}^{+} = 0.7$ V, $V_{te}^{-} = -0.9$ V, $V_{te}^{0} = 10$ mV, $V_{out}^{+} = 2$ V, $V_{th} = 9$ mV, $\tau_r = 3$ ms, $\tau_s = 50$ μs, $\tau_{out} = 1.5$ ms.

Fig. 5 shows the process of synaptic weight adaptation to recognized patterns. The shade of grey corresponds to the state variable value of the corresponding memristor: the darker, the greater the conductivity; the lighter, the less. At the initial moment of time, all weights are initialized with random values, and gradually change during the network operation. From about the 1200th era, patterns began to be seen, the recognition of which was trained by the network: the information was memorized by the neural network. The duration of one epoch is 1.5 ms.

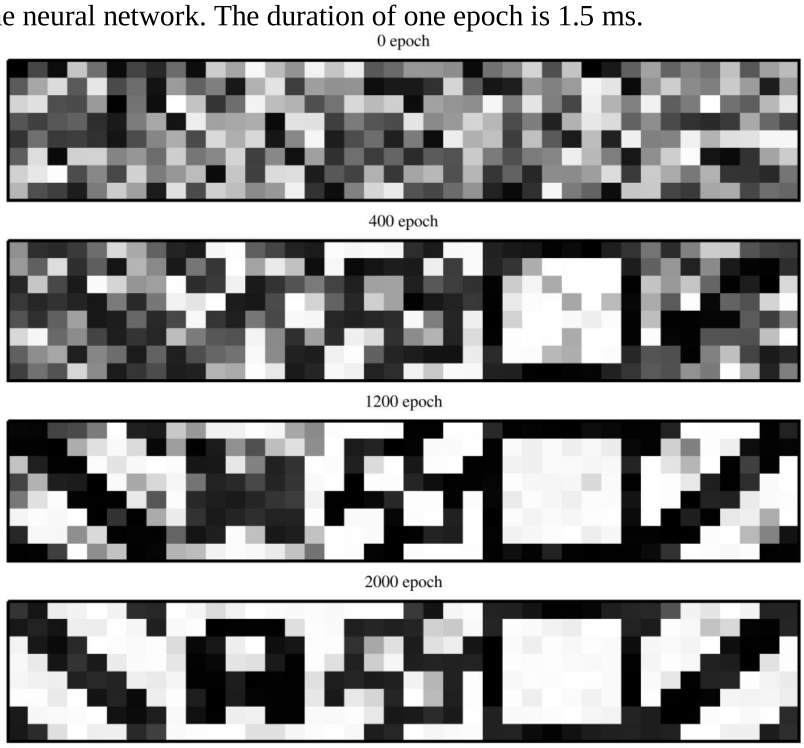

Figure 5: Change in synaptic weights in the process of training a neuromorphic network

The correspondence of the network weights to the patterns indicates that the network has successfully trained to recognize the given images. Due to the stochastic component in the memristor model, patterns can be distributed among neurons in different ways and the adaptation of weights can occur at different rates.

5. Conclusion

The work is devoted to simulating the operation of an analogue self-learning pulse neural network built on the basis of memristive elements in the image recognition mode. The simulation is carried out taking into account the stochastic properties of memristors. A variable-resistor thin-film memristor model based on an exponential dopant drift model, in which there is a term responsible for additive (Gaussian) noise, is considered. The ampere-voltage characteristics of the model are compared with experimental data on titanium oxide. The simulating of the operation for the network consisting of five neurons with 320 synapses, designed to recognize five different black-and-white images of 8 by 8 pixels, is performed. As a result the network has successfully trained to recognize the given patterns.

6. Acknowledgements

The work was performed with the support of RFBR grant No. 19-29-03051mk.

The studies were carried out using the resources of the Center for Shared Use of Scientific Equipment "Center for Processing and Storage of Scientific Data of the Far Eastern Branch of the Russian Academy of Sciences", funded by the Russian Federation represented by the Ministry of Science and Higher Education of the Russian Federation under project No. 075-15-2021-663.

7. References

[1] M. V. Egorchev, Y. V. Tiumentsev, Semi-empirical Neural Network Based Ap-proach to Modelling and Simulation of Controlled Dynamical Systems, Procedia Computer Science. 2018. Vol. 123. pp. 134–139. DOI: 10.1016/j.procs.2018.01.022.

[2] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza et al., Artificial brains. A million spiking-neuron integrated circuit with a scalable communication network and interface, Science. 2014 Aug 8; 345(6197):668-73. Epub 2014 Aug 7. PMID: 25104385. DOI: 10.1126/science.1254642.

[3] H.-S. P. Wong et al., Metal–oxide RRAM, Proceedings of the IEEE, 2012. Vol. 100, no. 6, pp. 1951–1970. DOI: 10.1109/JPROC.2012.2190369.

[4] J. J. Yang, D. B. Strukov, D. R. Stewart, Memristive devices for computing, Nature Nanotechnology, 2013. Vol. 8, no. 1, pp. 13–24. DOI: 10.1038/nnano.2012.240.

[5] C. Li, M. Hu, Y. Li, et al., Analogueue signal and image processing with large memristor crossbars, Nature Electronics. Vol. 1, no. 1, Dec. 2017, pp. 52–59. DOI: 10.1038/s41928-017-0002-z.

[6] A. Yu. Morozov, K. K. Abgaryan, D. L. Reviznikov, Issues of implementing neural network algorithms on memristor crossbars, Russian Microelectronics, 2020, Vol. 49, No. 8, pp. 568-573. DOI: 10.1134/S1063739720080053

[7] N. V. Agudov, A. V. Safonov, A. V. Krichiginet et al., Nonstationary distributions and relaxation times in a stochastic model of memristor, Journal of Statistical Mechanics: Theory and Experiment, 2020, 024003. DOI: 10.1088/1742-5468/ab684a

[8] A. N. Mikhaylov, E. G. Gryaznov, A. I. Belov et al., Field- and irradiation-induced phenomena in memristive nanomaterials, Phys. Status Solidi C - Current Topics in Solid State Physics, 2016; 13(10-12) pp. 870-881 DOI: 10.1002/pssc.201600083.

[9] D. O. Filatov, D. V. Vrzheshch, O. V. Tabakov et al., Noise-induced resistive switching in a memristor based on ZrO2(Y)/Ta2O5 stack, Journal of Statistical Mechanics: Theory and Experiment, 2019(12), 124026. DOI: 10.1088/1742-5468/ab5704.

[10] D. B. Strukov, G. S. Snider, D. R. Stewart, R. S. Williams, The missing memristor found, Nature. Vol. 453, no. 7191, 2008, p. 80. DOI: 10.1038/nature06932.

[11] J. J. Yang, M. D. Pickett, L. Xuema et al., Memristive switching mechanism for metal/oxide/metal nanodevices, Nature nanotechnology. Vol. 3, no. 7, 2008, p. 429. DOI: 10.1038/nnano.2008.160.

[12] Y. N. Joglekar, S. J. Wolf, The elusive memristor: properties of basic electrical circuits, European Journal of Physics. Vol. 30, no. 4, 2009, p. 661. DOI: 10.1088/0143-0807/30/4/001.

[13] V. A. Vasil'ev, P. S. Chernov, Mathematical modeling of memristor in the presence of noise, Matem. Mod., 2014. 26:1, pp. 122–132. (in Russian)

[14] A. Yu. Morozov, K. K. Abgaryan, D. L. Reviznikov, Mathematical model of a neuromorphic network based on memristive elements, Chaos, Solitons & Fractals, 2021, Vol. 143, 110548. DOI: 10.1016/j.chaos.2020.110548

[15] P. Diehl, M. Cook, Unsupervised learning of digit recognition using spike-timing-dependent plasticity, Frontiers in Computational Neuroscience, 2015. Vol. 9, art. 99, p. 9. DOI: 10.3389/fncom.2015.00099.
[16] S. Ambrogio et al. Neuromorphic Learning and Recognition With One-Transistor-One-Resistor Synapses and Bistable Metal Oxide RRAM, IEEE Transactions on Electron Devices, 2016. Vol. 63, no. 4, pp. 1508–1515. DOI: 10.1109/TED.2016.2526647.
[17] Y. Guo, H. Wu, B. Gao, H. Qian, Unsupervised Learning on Resistive Memory Array Based Spiking Neural Networks, Front Neurosci, 2019. Vol. 13, art. 812. DOI: 10.3389/fnins.2019.00812.
[18] V. Milo et al., Resistive switching synapses for unsupervised learning in feed-forward and recurrent neural networks, 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, 2018, pp. 1–5. DOI: 10.1109/ISCAS.2018.8351824.
[19] G. Zheng, S. P. Mohanty, E. Kougianos, O. Okobiah, Polynomial Metamodel integrated Verilog-AMS for memristor-based mixed-signal system design, 2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS), Columbus, OH, 2013, pp. 916-919. DOI: 10.1109/MWSCAS.2013.6674799.
[20] D. B. Strukov, G. S. Snider, D. R. Stewart, R. S. Williams, The missing memristor found, Nature. Vol. 453, no. 7191, 2008, p. 80. DOI: 10.1038/nature06932.
[21] A. Yu. Morozov, K. K. Abgaryan, D. L. Reviznikov Mathematical modeling of a self-learning neuromorphic network based on nanosized memristive elements with 1T1R crossbar architecture, Izvestiya Vysshikh Uchebnykh Zavedenii. Materialy Elektronnoi Tekhnik. 2020;23(3):186-195. (in Russian)