Using simulation to define the tolerances for the information and physical parameters of memristors-based artificial neural networks

S N Danilin, S A Shchanikov, I A Bordanov and A D Zuev

Vladimir State University named after Alexander and Nickolay Stoletovs, 87, Gorkogo Street, Vladimir, 600000, Russia

E-mail: seach@inbox.ru

Abstract. The article covers a solution to a problem of defining the tolerances of information and physical parameters of components of artificial neural networks (ANNs), which are implemented as hardware through the application of nanoscale electronic components with memristive properties (memristors). The developed method foundation is a system approach to the memristors-based ANN (ANNM) design, whereby the ANNMs should be studied as united physical and informational objects. When the ANNM is produced and operated, the errors of its components’ physical parameters provoke information parameter errors. To define the tolerated errors (tolerances), a simulation methodology is used. The potential of the developed method is illustrated through the process of defining tolerances for the synaptic weights and neural biases of a two-layer feed forward ANNM.

1. Introduction
The present-day artificial neural networks (ANN) are highly complex, yet they keep getting more sophisticated. The attained level of neural network theory [1] suggests that the analytical methods of the ANN synthesis are impossible to develop due to non-formalizable, multidimensional, nonlinear and probabilistic nature of the problems being solved, destabilizing factors, physical and information processes within ANN.

Currently, the most commonly used technique of implementing the ANNs in the solution of practical problems (as part of control systems, functional and diagnostic monitoring, decision support systems, computer vision, etc. [2, 3]) is their software and hardware emulation on von Neumann-architecture computers. A less commonly used technique of ANN implementing consists in the application of programmable logic devices (PLD) or specialized neuroprocessors. The choice of a particular method depends on multiple factors, such as the task complexity, required quality (accuracy, fault-tolerance, reliability, performance), project budget and so on.

A common feature of these methods is that the basic operations typical for the ANNs (weighted summation, bias, neuron activation) are performed by digital electronic components, in which the performance is always limited by the precision of an arithmetic logic unit, clock speed, and access rate to the instruction and data memory.

The analysis of the published scientific and technical research shows that the theory of the memristors-based ANN (ANNM) design, production and operation is at the early stage of development [4]. One of the paramount tasks of this theory is to develop engineering design methods and algorithms
for the ANNMs, and to define and ensure the required quality parameters of their performance [5]. The solution of this task will allow for the hardware implementation of the information, communication and radio systems algorithms [6, 7] for high-performance computing tools.

2. Approach

Russian scientists have made real progress regarding the development of production processes of both organic [8] and inorganic [9] memristors with the given properties and stable behavior. However, due to the physical properties of memristors, the materials from which they are produced, as well as the effects on which the resistive switch is based [10], there are errors in their operation. To ensure the required quality parameters of an ANNM (accuracy, fault-tolerance, reliability and performance), these errors should be kept within tolerance.

The approach to the ANNM engineering design, developed and applied by the authors, is based on the simulation methodology and allows for creation of components of their digital twins [11, 12].

Let us consider the potential of the developed approach through a method of defining the operation tolerances of a two-layer feed forward neural network, the neural synapses of which are implemented as hardware as a memristors bridge, consisting of the following:

- A simulation model of the ANNM with the specified parameters is created (architecture, structure (number of neurons and layers), learning algorithm, etc.).
- The ANNM performance metric is selected and the allowable value of error when solving a specific practical problem $X_a$ is specified.
- The ANNM is trained in accordance with the selected learning algorithm to achieve the best result according to the established metric $X_w (X_w << X_a)$.
- The errors of parameters of the ANNM components are simulated through a simultaneous change of their values $\pm \Delta m\%$ for all neurons.
- At each iteration, the current value of the metric $X_{i,j,k,l,f}$ is recorded.
- Using the formula (1), a relative metric of ANNs performance $X_{i,j,k,l,f}$ is calculated. If $K$ is greater than or equal to 0, we turn back to the previous item, otherwise we proceed to the next one:

$$K_{i,j,k,l,f} = 1 - \frac{X_{i,j,k,l,f} - X_w}{X_a - X_w}.$$ (1)

where $X_a$ is an allowable value (tolerance) of the ANNM performance metric; $X_w$ is a value of the ANNM performance metric achieved while training; $X_{i,j,k,l,f}$ is a value of the ANNM performance metric at parameter sweep of the $i$-th structure, $j$-th neuron, $k$-th component and $l$-th characteristics of the input data, $f$-th noise parameter of a reference value.

- A tolerance for the information and physical parameters of the ANNM components is defined.

Therefore, by changing the $\Delta m\%$ value in a cycle, the ANNM is driven to a state of failure ($K < 0$), and the tolerance of parameters of the neurons, in this case, will be equal to the previous value of $m\%$ when the network was still functioning.

3. Experiment

To determine the limits of the margin tolerance of the information parameters (in this case, the weights $w$ of neural synapses), we synthesize a model of a two-layer feed forward ANN (a multilayer perceptron). Let us consider a solution to the XOR problem as an example. The model parameters are the following: number of input neurons – 2, number of neurons of the first (hidden) layer – 2, number of neurons of the second (output) layer – 1. The functions of neuron activation are tangential. The mathematical model of the ANN is as follows:

$$\hat{y} = f \left( \sum_{j=1}^{N_{\text{hidden}}} w_{\text{out},j} \cdot f \left( \sum_{i=1}^{N_{\text{input}}} w_{i,j} \cdot x + b_{j} \right) + b_{\text{out}} \right),$$ (2)
where $\hat{y}$ is the ANN output data;

$j$ is a number of the hidden layer neuron ($N_{\text{hidden}} = 2$);

$i$ is a number of the input layer neuron ($N_{\text{input}} = 2$);

$w_{\text{out}}$ is an array of weights of the output neuron synapses;

$w$ are the synapses weights of the hidden layer neurons;

$x$ is the ANN input data;

$b$, $b_{\text{out}}$ are the hidden and output layer biases.

Let us assess the error of the ANN performance using a criterion of the mean sum squared absolute errors $MSE$ (functions of secondary optimization):

$$MSE = \frac{1}{H} \sum_{h=1}^{H} (y_h - \hat{y}_h)^2,$$

where $H$ is a total number of components of a training or a test set;

$y_h$ is the XOR function value for $h$-th value of the sampling;

$\hat{y}$ is the ANN output data value for $h$-th value of the sampling.

Upon setup, the ANN performance with respect to the $MSE$ was $X_o = 0.0007$. To define the margin of error for the hardware physical parameters, we synthesize an ANNM model (Figure 1).

Schematically, a synapse is a memristors bridge, described in the publications [13, 14].

Weight $W$ of a neural synapse is formed as follows [14]:

$$W = \frac{v_{\text{OUT}}}{v_{\text{IN}}} = \frac{M_2}{M_1 + M_2} - \frac{M_4}{M_3 + M_4},$$

where $v_{\text{IN}}$ is an input voltage amplitude of the bridge;

$v_{\text{OUT}}$ is an output voltage amplitude of the bridge;

$M_1, M_2, M_3, M_4$ are memristancies of four memristors of the bridge.

The memristor mathematical model is implemented according to the work [15].

![Figure 1](image_url)

**Figure 1.** A model of the ANNM neuron in the LTSpice software for defining the margin tolerance for values of the physical parameters (memristancies).

Let us experimentally define the tolerances for errors of the ANNM information and physical parameters in accordance with the following algorithm:

- We set the allowable value of the error $X_a = 2X_o$ with respect to the $MSE$ criterion.
• Using the ANN model, we calculate the weight tolerance using the worst-case approach by changing the weights of all the ANN synapses in a cycle by \( w \pm |w \Delta m| \), where \( \Delta m \) increases by 0.1; we record the current value of the performance error with respect to the MSE criterion \( X \) and calculate the value of the \( K \) metric. When \( K = 0 \), we stop the cycle and record the \( \Delta m \) value.

• Using the ANN synapse model, we calculate the memristancy tolerance of all components of the bridge.

4. Results and Discussion

Figure 2 shows the results of the ANN simulation with variations in the synaptic weights \( w \) and neural biases \( b \).

Consistent with the results presented in Figure 2, the allowable errors of the information parameters (weights and biases) for the present ANN are contained within \([-11.9\%; +7.7\%]\) range. When the values of the synaptic errors are outside the tolerance limits, the \( K \) assumes a negative value.

Using the simulation model of the ANNM neural synapse, let us define the tolerance for the information parameters error. Let us enter the results of the experiment into table 1. The ANNM performance will be out of tolerance, should the memristancy value for these memristors change by values greater than those indicated in table 1 for each synapse, in the process of production or operation during destabilizing influence.

| Synapse | Memristancy, Ohm |
|---------|------------------|
| \( W_{1.1} = W_{2.2} = W_{\text{out},1} = W_{\text{out},2} = 0.85 \) | \( W_{1.1} + \Delta m \) | \( +\Delta W_{1.1} \) | \( W_{1.1} - \Delta m \) | \( -\Delta W_{1.1} \) |
| \( M_1 \) | 37329.6 | 38648.6 | 1319 (+3.5%) | 35291.6 | 2038 (-5.45%) |
| \( M_2 \) | 3070.4 | 1751.4 | 1319 (-42.9%) | 5108.4 | 2038 (+66.4%) |
| \( M_3 \) | 3070.4 | 1751.4 | 1319 (-42.9%) | 5108.4 | 2038 (+66.4%) |
| \( M_4 \) | 37329.6 | 38648.6 | 1319 (+3.5%) | 35291.6 | 2038 (-5.45%) |

| Synapse | \( W_{1.2} = W_{2.1} = 0.75 \) | \( W_{1.1} + \Delta m \) | \( +\Delta W_{1.1} \) | \( W_{1.1} - \Delta m \) | \( -\Delta W_{1.1} \) |
|---------|------------------|------------------|------------------|------------------|------------------|
| \( M_1 \) | 5050.0 | 3883.3 | 1166.7 (-23.1%) | 6852.7 | 1802.7 (+35.7%) |
| \( M_2 \) | 35350.0 | 36516.7 | 1166.7 (+3.3%) | 33547.3 | 1802.7 (-5.1%) |
| \( M_3 \) | 35350.0 | 36516.7 | 1166.7 (+3.3%) | 33547.3 | 1802.7 (-5.1%) |
| \( M_4 \) | 5050.0 | 3883.3 | 1166.7 (-23.1%) | 6852.7 | 1802.7 (+35.7%) |

| Bias | \( B_1 = B_2 = B_{\text{out}} = 0.47 \) | \( W_{1.1} + \Delta m \) | \( +\Delta W_{1.1} \) | \( W_{1.1} - \Delta m \) | \( -\Delta W_{1.1} \) |
|------|------------------|------------------|------------------|------------------|------------------|
| \( M_1 \) | 10705.9 | 9974.6 | 731.3 (-6.8%) | 11835.5 | 1129.6 (+10.5%) |
| \( M_2 \) | 29694.0 | 30425.3 | 731.3 (+2.5%) | 28564.4 | 1129.6 (-3.8%) |
| \( M_3 \) | 29694.0 | 30425.3 | 731.3 (+2.5%) | 28564.4 | 1129.6 (-3.8%) |
| \( M_4 \) | 10705.9 | 9974.6 | 731.3 (-6.8%) | 11835.5 | 1129.6 (+10.5%) |

5. Conclusion

• The authors have proposed and justified the application of the methodology of the system approach and simulation of information processes for the engineering design of the information processing ANNM implemented based on various types of memristors.
Conforming to the applied approach, the ANNM simulation methods have been developed.

Simulation models for different levels of ANNMs (systems, subsystems, members, engineering elements) have been developed.

Using a test degree of the ANNM complexity as a case study, the authors have demonstrated the possibility to numerically define production tolerances for parameters of the ANNM memristors.

All the developed components of the ANNM engineering design were implemented as a software package in APL Python.

The results of this work suggest the long-term benefits of the proposed information technologies for design and research of ANNMs of commercial degree of complexity, free structure and purpose.

![Graph of the ANN performance vs. relative error of the weights w of neural synapses](image)

**Figure 2.** A graph of the ANN performance vs. relative error of the weights $w$ of neural synapses: a – with respect to the $K$ metric; b – with respect to the $MSE$ criterion.

**Acknowledgments**

This study was conducted with the support of the Russian Foundation for Basic Research, grants No. 18-38-00592.

**References**

[1] Galushkin A I 2007 *Neural Networks Theory* (Heidelberg: Springer-Verlag Berlin) p 396

[2] Avedyan E D, Galushkin A I and Pantiukhin D V 2014 Neural network technologies for information systems security *7th IEEE Symp. on Computational Intelligence for Security and Defense Applications* (New York: Springer US)
[3] Zhang Y D et al 2019 Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation Multimedia Tools and Applications 78(3) 3613-32
[4] Lanza M et al 2019 Recommended Methods to Study Resistive Switching Devices Advanced Electronic Materials (Weinheim: Wiley-VCH)
[5] Galushkin A I 2014 Neural networks realizations using memristors 7th Int. Conf. on Engineering and Telecommunication (Moscow: Institute of Electrical and Electronics Engineers) pp 77-81
[6] Zhiganov S N, Zhiganova E A and Mikheev K V 2019 Development of the Method for Constructing Trajectories of Air Objects 2018 Int. Multi-Conf. on Industrial Engineering and Modern Technologies (Vladivostok: Institute of Electrical and Electronics Engineers)
[7] Thang V V, Pantiukhin D V and Nazarov A N 2016 FLDS: Fast outlier detection based on local density score 2016 Int. Conf. on Engineering and Telecommunication (Moscow: Institute of Electrical and Electronics Engineers) pp 137-41
[8] Lapkin D A, Emelyanov A V, Demin V A, Erokhin V V, Feigin L A, Kashkarov P K and Kovalchuk M V 2018 Polyaniline-based memristive microdevice with high switching rate and endurance Applied Physics Letters 112(4) (Melville: American Institute of Physics)
[9] Antonov I N, Belov A I, Mikhaylov A N, Morozov O A and Ovchinnikov P E 2018 Formation of Weighting Coefficients in an Artificial Neural Network Based on the Memristive Effect in Metal–Oxide–Metal Nanostructures Journal of Communications Technology and Electronics 63(8) 950-57
[10] Chua L O 2018 Five non-volatile memristor enigmas solved Applied Physics A 124 563-606
[11] Danilin S N, Shchanikov S A and Sakulin A E 2017 Algorithm for Determining Optimum Operation Tolerances of Memristor-Based Artificial Neural Networks 2017 IVth Int. Conf. on Engineering and Telecommunication (Vladivostok: Institute of Electrical and Electronics Engineers) pp 140-44
[12] Danilin S N, Shchanikov S A and Panteleev S V 2016 Determining Operation Tolerances of Memristor-Based Artificial Neural Networks 2016 Int. Conf. on Engineering and Telecommunication (Moscow: Institute of Electrical and Electronics Engineers) pp 34-8
[13] Adhikari S P et al 2012 Memristor Bridge Synapse-Based Neural Network and Its Learning IEEE Transactions on neural networks and learning systems 23(9)
[14] Yang C, Kim H, Adhikari S P and Chua L O 2017 A Circuit-Based Neural Network with Hybrid Learning of Backpropagation and Random Weight Change Algorithms Sensors 17(1) (Cairo: Hindawi Publishing Corporation)
[15] Strukov D B 2008 The missing memristor found Nature 453(7191) 80-3