# Precise Modeling of Emerging Electronic Structures by Artificial Neural Networks

Josef Dobeš, Ladislav Pospíšil, and Abhimanyu Yadav

***Abstract*—Nowadays, there are many emerging electronic structures for which their nonlinear models for computer-aided design are necessary, especially for the ones from the areas of nanoelectronics and microwave techniques. However, for such structures, sufficiently accurate analytic models are mostly unavailable. This is partially caused by the fact that the physical principles of the element operation are sometimes not fully clear (especially for quantum devices), and also by bizarre characteristics of some of the elements (typically with irregularities and a hysteresis in parts of characteristics, or by negative differential conductances that are typical for the microwave transistors). In such cases, models based on artificial neural networks are necessary and useful for these elements. Majority of the elements can be characterized with a single artificial neural network. However, for certain kinds of elements, a cooperation of more artificial neural networks is necessary. This case is described in the paper first, where the Pt−$TiO_{2-x}$−Pt memristor characteristic with an extraordinary (but typical) hysteresis is approximated by a set of cooperative artificial neural networks, as a single network is unable to characterize this unconventional element. Second, an ability of the artificial neural networks for modeling the negative differential conductance is demonstrated by characterizing the 110GHz pseudomorphic high electron mobility transistor (pHEMT). Moreover, a semiautomatic selection of an optimal structure of the networks (both numbers of hidden layers and the numbers of the elements in the layers) is also suggested.**

***Index Terms*—artificial neural networks, multilayer perceptron, device characterization, memristive system, pHEMT.**

## I. Introduction

A MATHEMATICAL description of the memristive elements is a very complicated task due to an extraordinary hysteresis and several irregularities in some parts of their characteristics. It can be easily seen in Fig. 1, where a set of measured points for a Pt − $TiO_{2-x}$ − Pt memristor is shown [1]. (The element is called "memristor" in [1]; however, it is rather a memristive system because the characteristic does not have odd symmetry [2], [3].) There are various analytic models of memristors [2], but their precision is naturally limited due to the simplicity of the functions. Even single – both exclusive and corrective – artificial neural networks (of the multilayer perceptron type shown in Fig. 2) are unable to characterize this element due to the fact it is neither $i = f(v)$ nor $v = f(i)$ function. For this reason, the element has to be approximated by more *cooperative* neural networks, some of them should be of $i = f(v)$ and other ones of $i = f(v)$ types.

Manuscript received July 23, 2012; revised August 10, 2012. This paper has been supported by the Grant Agency of the Czech Republic, grant No. P102/10/1614, and by the Czech Technical University in Prague internal research grants Nos. SGS 11/160/OHK3/3T/13, SGS 10/286/OHK3/3T/13, and SGS 12/151/OHK3/2T/13.

The authors are with the Department of Radio Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha 6, Czech Republic.

E-mails: {dobes, pospisl, yadavabh}@fel.cvut.cz

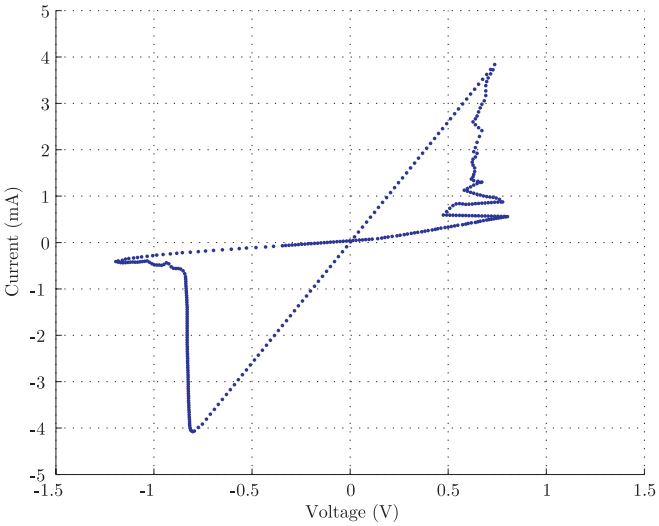


Fig. 1. Measured characteristic of the memristive system.

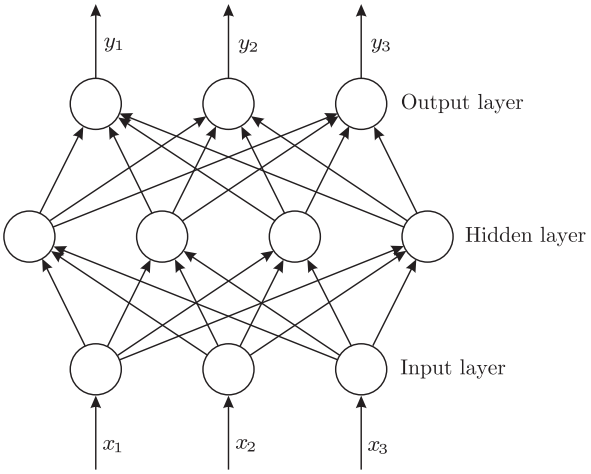


Fig. 2. Multilayer perceptron (MLP) structure.

Generally, for the pHEMT device, it is important to emphasize that the nonlinear dependence of the drain current on the gate-drain and gate-source voltages affects its overall behavior, i.e., also the hundred-gigahertz part of the model is strongly affected by the transistor characteristics [4].

Further, for modeling the pHEMT devices, the two regions in the characteristics are critical regarding the precision:

- The low-current part (for a subthreshold voltage, or slightly over the threshold voltage), where the dependence of the threshold voltage on the drain-source voltage becomes extremely important (and frequently inaccurately defined by standard analytic models).
- The high-current part, where the negative output conductance becomes substantial (and again inaccurately defined in most of the analytic models).

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
SYSTEMATIC SEARCHING FOR THE OPTIMAL NUMBER OF ELEMENTS OF THE 5-LAYER STRUCTURE FOR REGION #1

| Type of network $i = f(v)$ 5-layer | Relative deviation after 500 training epochs rms (%) |
|---|---|
| MLP-1-2-2-2-1 | 2.67 |
| MLP-1-3-2-2-1 | 0.93 |
| MLP-1-4-2-2-1 | 0.87 |
| MLP-1-2-3-2-1 | 0.50 |
| MLP-1-2-4-2-1 | 0.65 |
| MLP-1-2-2-3-1 | 1.15 |
| MLP-1-2-2-4-1 | 0.79 |

TABLE II
SYSTEMATIC SEARCHING FOR THE OPTIMAL NUMBER OF ELEMENTS OF THE 5-LAYER STRUCTURE FOR REGION #2

| Type of network $i = f(v)$ 5-layer | Relative deviation after 500 training epochs rms (%) |
|---|---|
| MLP-1-2-2-2-1 | 3.88 |
| MLP-1-2-3-2-1 | 3.91 |
| MLP-1-2-4-2-1 | 1.75 |
| MLP-1-2-2-3-1 | 4.02 |
| MLP-1-2-2-4-1 | 1.05 |
| MLP-1-3-2-3-1 | 2.25 |
| MLP-1-4-2-2-1 | 1.51 |

TABLE III
SYSTEMATIC SEARCHING FOR THE OPTIMAL NUMBER OF ELEMENTS OF THE 5-LAYER STRUCTURE FOR REGIONS #3 & #4

| Type of network $i = f(v)$ 5-layer | Relative deviation after 500 training epochs rms (%) |
|---|---|
| MLP-1-3-3-2-1 | 12.25 |
| MLP-1-3-4-2-1 | 9.86 |
| MLP-1-3-5-2-1 | 3.16 |
| MLP-1-3-6-2-1 | 2.95 |
| MLP-1-3-7-2-1 | 2.80 |
| MLP-1-3-8-2-1 | 3.05 |

TABLE IV
SYSTEMATIC SEARCHING FOR THE OPTIMAL NUMBER OF ELEMENTS OF THE 6-LAYER STRUCTURE FOR REGION #5

| Type of network $v = f(i)$ 6-layer | Relative deviation after 500 training epochs rms (%) |
|---|---|
| MLP-1-3-3-3-2-1 | 13.22 |
| MLP-1-3-4-3-2-1 | 15.28 |
| MLP-1-3-4-4-2-1 | 11.80 |
| MLP-1-3-5-4-2-1 | 8.05 |
| MLP-1-3-6-4-2-1 | 7.83 |
| MLP-1-3-6-5-2-1 | 8.75 |
| MLP-1-3-7-6-2-1 | 7.94 |

However, the two above phenomena can be easily and precisely characterized by multilayer artificial neural networks.

## II. A STRATEGY OF USING THE COOPERATIVE ARTIFICIAL NEURAL NETWORKS FOR MODELING THE MEMRISTIVE SYSTEMS

As the single networks, the MLP (multilayer perceptron) ones were used. For example, for the MLP-3-4-3 network in Fig. 2, we have three input $\boldsymbol{x} = (x_1, x_2, x_3)$ and three output $\boldsymbol{y} = (y_1, y_2, y_3)$ variables, which are determined by the formula

$$\boldsymbol{y} = F^3\left[\boldsymbol{W^{32}} F^2\left(\boldsymbol{W^{21}}\boldsymbol{x} - w_0^2\right) - w_0^3\right], \qquad (1)$$

where the single superscript marks the layer number, and the double one expresses the relation of the two layers. The scalars $w_0^2$ and $w_0^3$ are threshold levels of neurons ($w_0$) of the second (hidden) and third (output) layers, respectively. The matrices $\boldsymbol{W^{32}}$ and $\boldsymbol{W^{21}}$ contain synaptic weights [5] between the second and third layers, and first and second ones, respectively; and the operators $F^3$ and $F^2$ symbolize the nonlinear functions in the third and second layer, respectively.

The result of using the four cooperative neural networks is shown in Fig. 3. (For some more bizarre types of the memristive systems, the regions #3&4 in Fig. 3 cannot be represented by one neural network, but by two ones.) Any of the four regions (#1, #2, #3&4, and #5) were modeled by a single artificial neural network; however, they must cooperate because they share boundary points. The comparison of some measured points (dots) with the model (solid lines) in Fig. 3 demonstrates the successfulness of the method. Moreover, for each neural network, an optimal structure of its hidden layers was also investigated, and the results are shown in Tables I through IV. For any network, there exists an optimal structure, it is not the most complicated case, and there are more solutions near the optimum. It is clear that the root-mean-square error rms can be of a percentage order even for the approximation of the most problematic part of the memristive-system characteristic.

## III. USING THE SINGLE MULTILAYER ARTIFICIAL NEURAL NETWORK FOR MODELING THE PSEUDOMORPHIC HIGH ELECTRON MOBILITY TRANSISTOR (PHEMT)

In order to demonstrate the performance of the identification procedure, measured multibias data of the transistor of a GaAs-pHEMT type with a gate length $0.25$ μm has been used [6], [7].

After numerical experiments, we have found that a five-layer network is necessary for sufficiently precise characterization of this 110GHz transistor. A typical result for the MLP-2-9-13-8-1 artificial neural network is shown in Fig. 4 (again, (1) was used). The network yields very precise results in practically all the parts of the characteristics.

Moreover, similarly to the investigations for the memristive system, we have searched for an optimal structure of the network with respect to its accuracy. Some of the results are shown in Table V. The best results – for both root-mean-square error rms (2) and maximum-absolute-value-deviation $\delta_{\max}$ (3) – were obtained for the MLP-2-10-12-8-1 structure. However, there are more structures that enable similar errors, e.g., MLP-2-10-11-8-1 or MLP-2-9-12-8-1. That means there are no critical demands for finding just one optimal structure – more ones can offer comparable accuracy.

Furthermore, from Table V, it can also be seen that the most complicated network MLP-2-10-13-8-1 has given the worst result. It is not an exceptional feature – the experiments have shown there exists a region of relatively not complicated networks that give the most accurate results.

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

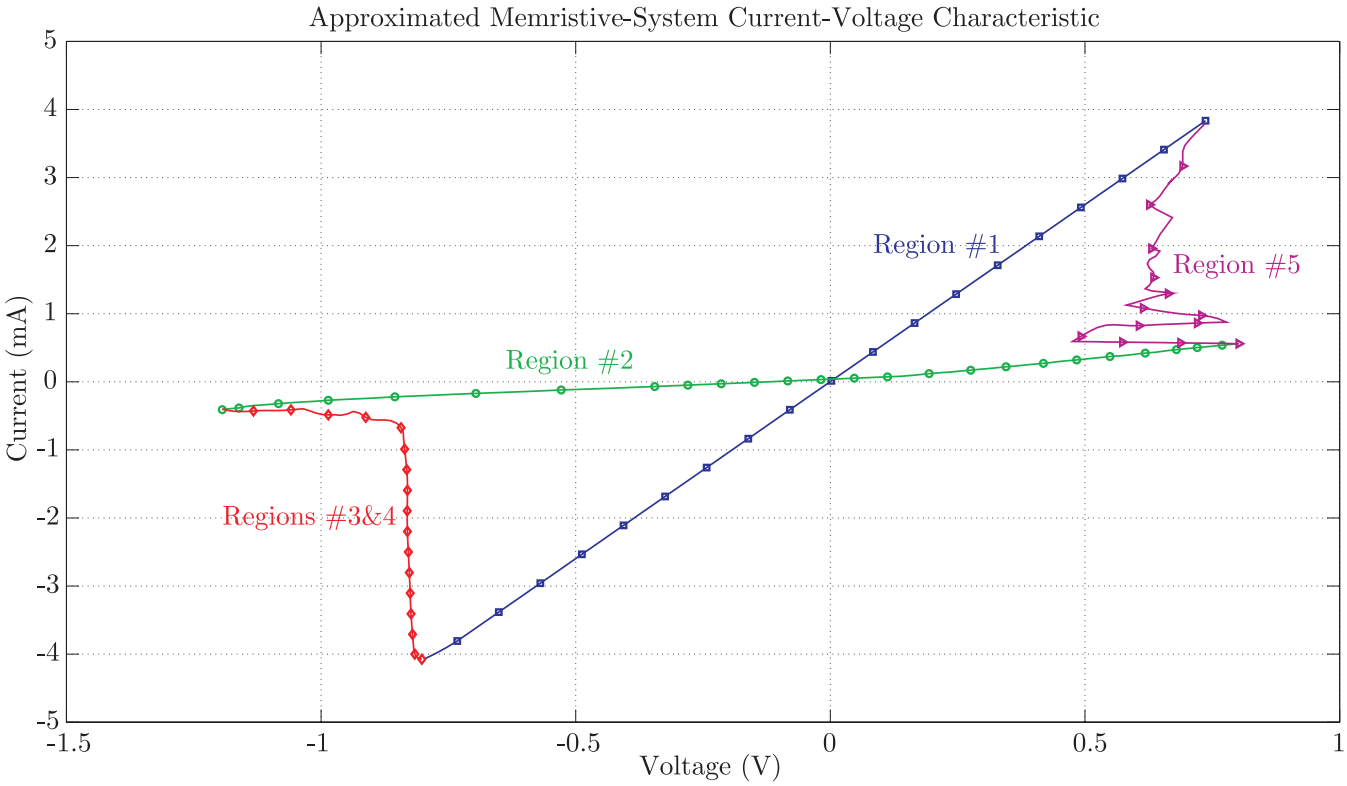


Fig. 3. An approximation of the memristive system with a set of cooperative artificial neural networks, both $y = f(x)$ and $x = f(y)$ types of the networks are necessary because any single network is unable to characterize the element precisely.

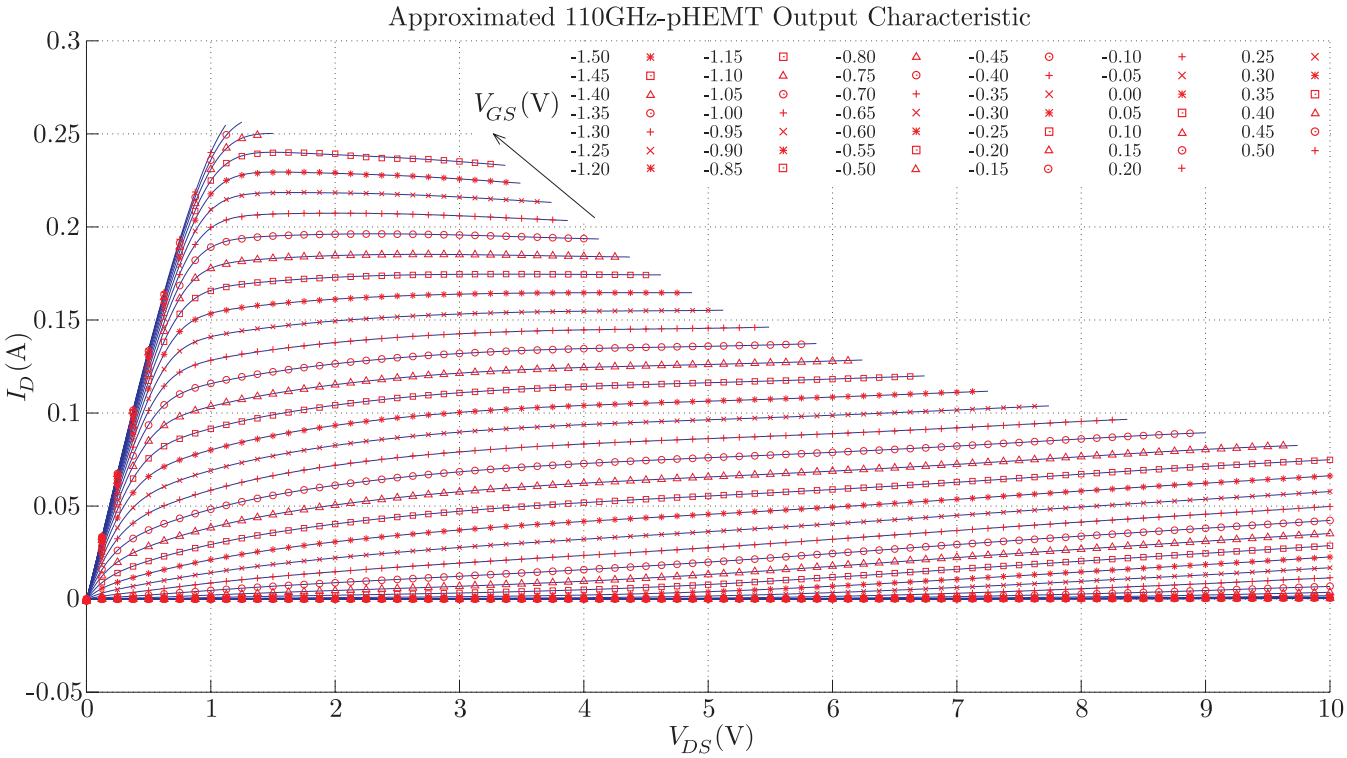


Fig. 4. An approximation of the 110GHz 0.25μm GaAs pHEMT output characteristics by the single MLP-2-9-13-8-1 artificial neural network.

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE V
SYSTEMATIC SEARCHING FOR THE OPTIMAL NUMBER OF ELEMENTS OF THE 5-LAYER STRUCTURE FOR CHARACTERIZING PHEMT

| Type of network | rms (%) | $\delta_{\max}$ (%) |
|---|---|---|
| MLP-2-10-13-8-1 | 0.8568 | 71.0582 |
| MLP-2-10-12-8-1 | 0.2835 | 10.1848 |
| MLP-2-10-11-8-1 | 0.3406 | 12.0719 |
| MLP-2- 9-13-8-1 | 0.4203 | 17.7939 |
| MLP-2- 9-12-8-1 | 0.3116 | 13.3131 |
| MLP-2- 9-11-8-1 | 0.3565 | 22.811 |

In a summary, with characterizing pHEMTs by a convenient artificial neural network (which can be found relatively easily), we can obtain $\mathrm{rms}$ lesser than a percent and $\delta_{\max}$ of about 10 percent.

## IV. CONCLUSIONS

In the paper, a reliable and robust method for the characterization of emerging electronic elements by the artificial neural networks has been presented. For the elements with strong hysteresis and substantial irregularities, an original method has also been suggested based on more cooperative networks. A semiautomatic method for selecting the optimal structure of single neural networks has been demonstrated, which shows that the resulting error can be of a percentage order even for the elements with extraordinary hysteresis and irregularities. In this way, using the novel idea of cooperative neural networks promises to be a sufficient modeling tool for various structures of nanoelectronics. Furthermore, for modeling pHEMTs and similar microwave transistors, only a single network can be sufficiently precise, which has been proven by a number of experiments, and the optimal structure of the network can be found relatively easily as well.

## APPENDIX

The root-mean-square error ($\mathrm{rms}$) and maximum-absolute-value-deviation ($\delta_{\max}$) computed for the results in Tables I–V are defined in the following natural way:

$$\mathrm{rms} = \sqrt{\frac{\sum_{i=1}^{n_p}\left(\frac{y_i^{(\mathrm{ident})} - y_i^{(\mathrm{meas})}}{y_i^{(\mathrm{meas})}}\right)^2}{n_p}} \times 100\,\%, \qquad (2)$$

$$\delta_{\max} = \max_{i=1,\ldots,n_p}\left|\frac{y_i^{(\mathrm{ident})} - y_i^{(\mathrm{meas})}}{y_i^{(\mathrm{meas})}}\right| \times 100\,\%, \qquad (3)$$

respectively, where $y_i^{(\mathrm{ident})}$ and $y_i^{(\mathrm{meas})}$ are the identified and measured values, respectively, and $n_p$ is the number of all points.

## ACKNOWLEDGMENT

Authors thank to prof. Santarelli and his colleagues from the University of Bologna for providing the measured data set.

## REFERENCES

[1] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, pp. 80–83, May 2008.
[2] Z. Biolek, D. Biolek, and V. Biolková, "SPICE model of memristor with nonlinear dopant drift," *Radioengineering*, vol. 18, no. 2, part II, pp. 210–214, June 2009.
[3] ——, "Analytical solution of circuits employing voltage- and current-excited memristors," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. PP, no. 99, pp. 1–10, 2012 (early access article).
[4] J. Dobeš and M. Grábner, "Novel HEMT models with improved higher-order derivatives and extracting their parameters using multibias S-parameters," in *Proc. of IEEE Compound Semiconductor IC Symp.*, Waikoloa Village, Big Island, Hawaii, Oct. 2011, pp. 189–192.
[5] Q. J. Zhang and K. C. Gupta, *Neural networks for RF and microwave design*. Boston: Artech House, 2000.
[6] A. Raffo, J. Lonac, D. Resca, S. Monaco, A. Santarelli, and G. Vannini, "Comparison of electron device models based on operation-specific metrics," in *Proc. of Asia-Pacific Microwave Conf.*, Suzhou, China, Dec. 2005, pp. 2411–2414.
[7] A. Raffo, A. Santarelli, P. A. Traverso, G. Vannini, F. Palomba, F. Scappaviva, M. Pagani, and F. Filicori, "Accurate pHEMT nonlinear modeling in the presence of low-frequency dispersive effects," *IEEE Trans. Microw. Theory Tech.*, vol. 53, pp. 3449–3459, Nov. 2005.

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)