Two Neural Approaches for Small-Signal Modelling of GaAs HEMTs

Zlatica Marinković, Giovanni Crupi, Alina Caddemi, Vera Marković

Abstract — Focus of this paper is on the neural approach in small-signal modelling of GaAs HEMTs. Two modelling approaches based on artificial neural networks are discussed and compared. The first approach is completely based on artificial neural networks, while the second is a hybrid approach putting together artificial neural networks and an equivalent circuit representation of a microwave transistor. Both models consider the device gate width and therefore both are scalable. Results of modelling of three different AlGaAs/GaAs HEMTs in a wide range of operating bias conditions using the considered approaches are given. Different modelling aspects are discussed. A special attention is paid to the model development procedure and accuracy of the models.

Index Terms—Artificial neural networks, HEMT, gate width, small-signal modelling

I. INTRODUCTION

RELIABLE and accurate models of microwave devices are necessary for the successful microwave circuit design. Therefore, developing of such models is an important and challenging task for researchers in the microwave community. Here the attention is paid to the modelling of the scattering (S-) parameters of microwave HEMTs (High Electron Mobility Transistors). The S-parameters of a microwave transistor depend on the operating bias condition as well as of the frequency. The most common modelling approach is to represent a device with a suitable equivalent circuit, whose ECPs are extracted from the measured values of the S-parameters, and to simulate the S-parameters in a microwave simulator. Although the ECP extraction is usually an ill-conditioned problem, there are efficient procedures for the ECP extraction [1]-[5], as the one that is used in this work [5]. The common feature for all equivalent circuit based models is that for each operating point it is necessary to repeat extraction of the ECPs, which can be quite time consuming if the ECPs should be extracted in a wide range of operating conditions. Alternative approach in modelling of the S-parameters is based on application of artificial neural networks (ANNs). In the last two decades artificial neural networks have found their place as an efficient tool for modelling of microwave devices [6], [7]. ANN models are usually extracted from the measured data directly, without need for detailed knowledge about device physics, therefore they can encounter all effects contributing to the device behaviour. As far as microwave transistors are concerned, ANNs have been exploited to model their small-signal, large-signal, and noise performance [8]-[12]. This paper is focused on multi-bias small-signal modelling of GaAs HEMTs differing in the gate width.

Recently, two different approaches based on the application of ANNs have been proposed to model scaled microwave FETs [11]-[12]. One approach is completely based on the neural networks, i.e. it is a black box model, while the other approach is a hybrid model, in which ANNs are used to predict parameters of an equivalent circuit representation. The aim of this paper is to compare both approaches in order to highlight advantages as well as disadvantages of both models in terms of modelling efficiency and accuracy. This paper is an extended version of the work presented in [13] giving more modelling details.

The paper is organized as follows. First, in Section II basic background on artificial neural networks is given. In Section III and IV the considered modelling approaches are presented. Section IV gives also the description of the analytical procedure used in the hybrid approach for the extraction of the equivalent circuit parameters (ECPs) from the measured scattering parameters of the modelled device. The modelling results obtained by using the two mentioned approaches on three AlGaAs/GaAs HEMT devices with different gate width are presented in Section V. The main conclusions are stated in Section VI.

II. ARTIFICIAL NEURAL NETWORKS

A widely used type of artificial neural networks is a standard multilayer perceptron (MLP) ANN shown in Fig. 1, [6]. This network consists of an input layer (layer 0), an output layer (layer NL), as well as several hidden layers. Input vectors are presented to the input layer and fed through the network that then yields the output vector. The l-th layer output is:

\[ Y_l = F(W_l Y_{l-1} + B_l) \]  \hspace{1cm} (1)

where \( Y_l \) and \( Y_{l-1} \) are outputs of l-th and (l-1)-th layer, respectively, \( W_l \) is a weight matrix between (l-1)-th and l-th layer.
th layer and $B_l$ is a bias matrix between $(l-1)$-th and $l$-th layer.

Fig. 1 MLP neural network

The function $F$ is an activation function of each neuron and, in our case, is linear for input and output layer and sigmoid for hidden layers:

$$F(u) = \frac{1}{1+e^{-u}}.$$  

The neural network learns relationship among sets of input-output data (training sets). First, input vectors are presented to the input neurons and output vectors are computed. These output vectors are then compared with desired values and errors are computed. Error derivatives are then calculated and summed up for each weight and bias until whole training set has been presented to the network. These error derivatives are then used to update the weights and biases for neurons in the model. The training process proceeds until errors are lower than the prescribed values or until the maximum number of epochs is reached (epoch is the whole training set processing). Once trained, the network provides fast response for various input vectors (even for those not included in the training set) without additional optimizations.

Usually, neural networks with different number of hidden neurons are trained, tested and after their comparison, the network with the best modeling results is chosen as the neural model.

III. BLACK-BOX SMALL-SIGNAL NEURAL MODEL OF MICROWAVE TRANSISTORS

The simplest type of ANN based models are the black-box models consisting of one or more MLP ANNs trained to predict dependence of the modelled parameters on the desired input conditions. As far as small-signal model of microwave FETs differing in the gate width is concerned, the modelled parameters are the S-parameters, while the inputs are: bias voltages (gate-to-source voltage, $V_{gs}$, and drain-to-source voltage, $V_{ds}$), frequency as well as the gate width, $W$ [11]. Therefore, there are four neurons in the input layer corresponding to the mentioned four input parameters. There are eight outputs in total (real and imaginary parts of four S-parameters). All eight outputs can be modelled by a single ANN, but in order to achieve better modelling accuracy it is recommended to train an ANN for each real and imaginary part of the S-parameters. The model is illustrated in Fig. 2. The ANNs used in the model are trained by using the S-parameters of the considered different gate width devices measured at a certain number of bias-points. Once the ANNs are trained properly, the S-parameters for any of the considered devices can be determined directly from the ANNs, by finding the response of the ANNs for the given input values.

IV. HYBRID SMALL-SIGNAL NEURAL MODEL OF MICROWAVE TRANSISTORS

As it was stated in Introduction, an equivalent circuit representation requires repeated extraction of the ECPs for each operating point. There are different extraction techniques, either based on optimizations or on analytical procedures. In order to avoid repeated extractions, one or more ANNs can be trained to predict dependence of the ECPs on the biases and the gate width [12]. For training of the ANNs the ECP values should be extracted from the measured S-parameters for certain number of different input conditions by using an extraction procedure. After the ANNs have been trained, the ECPs can be determined by finding the response of the ANNs. Using the predicted ECPs, the S-parameters are simulated from the equivalent circuit, as represented in Fig. 3.

IV. HYBRID SMALL-SIGNAL NEURAL MODEL OF MICROWAVE TRANSISTORS

As it was stated in Introduction, an equivalent circuit representation requires repeated extraction of the ECPs for each operating point. There are different extraction techniques, either based on optimizations or on analytical procedures. In order to avoid repeated extractions, one or more ANNs can be trained to predict dependence of the ECPs on the biases and the gate width [12]. For training of the ANNs the ECP values should be extracted from the measured S-parameters for certain number of different input conditions by using an extraction procedure. After the ANNs have been trained, the ECPs can be determined by finding the response of the ANNs. Using the predicted ECPs, the S-parameters are simulated from the equivalent circuit, as represented in Fig. 3.

This approach is a hybrid technique since the proposed model consists of ANNs, which are trained to predict the ECPs, and an equivalent circuit. The developed model can be used in the microwave circuit simulator as a library element for the considered scaled devices.
Equivalent circuit model considered in this work is shown in Fig. 4. The eight ECPs of the intrinsic circuit \((C_{gs}, R_{gs}, C_{gd}, R_{gd}, g_m, \tau, R_{ds}, C_{ds})\) are dependent on the biases, while the eight ECPs of the extrinsic circuit are assumed to be bias independent.

In [5] an analytical technique for the direct extraction of the ECPs of such circuit was proposed. The extrinsic ECPs are extracted analytically from the S-parameters under “cold pinch-off” condition. After removing their contributions from the data, the intrinsic ECPs are calculated straightforwardly at each bias point from the intrinsic admittance \((Y-)\) parameters.

Having in mind that the extrinsic ECPs are not bias dependent, the ANN modelling them has only one input neuron corresponding to the gate width, Fig. 5. The seven outputs of the network correspond to the seven extrinsic ECPs (it was assumed that \(C_{pg} = C_{pd}\)). The ANNs modelling the intrinsic ECPs have, beside the gate width, the bias voltages \(V_{gs}\) and \(V_{ds}\) as inputs; therefore each of them have three input neurons. There are eight output neurons of this ANN model, corresponding to eight intrinsic ECPs.

![Fig. 5. Hybrid model for the considered devices](image)

### V. MODELLING RESULTS

Both of the above mentioned approaches were applied to on-wafer AlGaAs/GaAs HEMTs with different gate width 100, 200, and 300 \(\mu\)m [11]-[13]. The models were developed from the S-parameters measured with an Agilent E8364A PNA (precision network analyzer). The measurements were done in 546 bias points.

#### A. Black-box model

As far as the black-box model is considered, the S-parameters were modelled by separate ANNs. For \(S_{11}\) and \(S_{22}\), a single ANN is used to model both real and imaginary parts of each parameter. But, for the two other parameters, \(S_{21}\) and \(S_{12}\), due to their more complicated behaviour, the real and imaginary parts of each parameter were modelled by separate ANNs. Therefore, the model consists of two two-output ANNs for \(S_{11}\) and \(S_{22}\) and four one-output ANNs for \(S_{21}\) and \(S_{12}\) as it is shown in Fig. 6. Each ANN was trained separately.

Details about the trained ANNs chosen for the final model are given in Fig. 5, where each ANN is described by four numbers \(N-H1-H2-M\). \(N\) is a number of neurons in the input layer, \(M\) is a number of neurons in the output layer, while \(H1\) and \(H2\) are numbers of neurons in the first and in the second hidden layer, respectively. More details about the ANN training can be found in [11].

#### B. Hybrid model

The starting point for the hybrid model of the considered devices was the equivalent circuit representation shown in Fig. 4. The ECPs were extracted analytically in a way described in [5].

The corresponding hybrid model with the information about the trained ANNs chosen for the final model (using the same notation as in the case of black-box model) is depicted in Fig. 6.

![Fig. 6. Developed black-box model](image)

It should be noted that in this case, \(g_m\) and \(\tau\) were modelled by separate one-output ANNs, while rest of the parameters were modelled by a single six-output ANN. More details on the model development are given in [12].

![Fig. 7. Developed hybrid model](image)
C. Model Comparison

Both models were implemented into the circuit simulator by implementing the mathematical expression describing the trained ANNs. After that, the S-parameters were simulated over the frequency range extending from 0.5 up to 50 GHz, at 546 different bias points (0 V < $V_{ds}$ < 2.5 V with 100 mV step and -1.5 V < $V_{gs}$ < 0 V with 75 mV step) and compared with the measured data, [13].

For the purpose of comparing the simulated and the measured S-parameters at each particular bias point, the percentage errors $E_{ij}$ were calculated as follows:

$$E_{ij} = \frac{1}{N_f} \sum_{f=1}^{N_f} \left( \frac{S_{ij}^{\text{MEAS}}(f) - S_{ij}^{\text{SIM}}(f)}{S_{ij}^{\text{MEAS}}(f)} \right)$$

where $N_f$ represents the number of frequency points per a bias point, which is 101 in the present case.

The total percentage error of the model, $E_{TOT}$, was obtained by averaging the evaluated errors of the four $S_{ij}$ of each device.

Since the modelling results of all three devices are very similar, for sake of space, the comparative analysis is focused on the device with 100 μm gate width. In Table I there are the percentage errors averaged for all considered bias points for the 100 μm device, for both neural models as well as for the analytical approach. It can be seen that the black-box model gives the smallest percentage errors, which can be explained by the fact that the ANNs were trained using the measured S-parameters without the constraint of the specific equivalent circuit topology and therefore all effects contributing to the device behaviour are included in the model. On the other side, as can be expected, the hybrid model gives slightly bigger percentage errors than the analytical one. This is due to the fact that the ANNs in the hybrid model are trained using the data obtained by the analytical approach.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>PERCENTAGE ERROR AVERAGED OVER THE FULL BIAS RANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 μm device</td>
<td>analytical</td>
</tr>
<tr>
<td>$E_{11}$[%]</td>
<td>11.0</td>
</tr>
<tr>
<td>$E_{21}$[%]</td>
<td>10.2</td>
</tr>
<tr>
<td>$E_{12}$[%]</td>
<td>10.2</td>
</tr>
<tr>
<td>$E_{22}$[%]</td>
<td>5.4</td>
</tr>
<tr>
<td>$E_{TOT}$[%]</td>
<td>9.2</td>
</tr>
</tbody>
</table>

One can see that the $S_{21}$ parameter in the case of the black-box model exhibits a significantly larger error. In order to explain this, Fig. 8 shows the bias dependence of the percentage error associated to $S_{21}$ for all three approaches. It can be observed that in the case of the black-box model the percentage error exhibits large values for the biases where $S_{21}$ is small (i.e., "cold" and "pinch-off" conditions) compared to its high values under typical operating conditions. This can be attributed to the fact that the neural approach is strongly based on its generalization capability, while in the analytical approach the equivalent circuit is extracted from the measurements at each specific bias point without accounting for the device behavior under the other bias points. In order to illustrate this more clearly, Fig. 9 illustrates the real and imaginary parts of $S_{21}$ obtained by the black-box model for two bias points: one referring to typical bias condition ($V_{gs} = -0.6$ V, $V_{ds} = 2.5$ V) and the other referring to "pinch-off" condition ($V_{gs} = -1.5$ V, $V_{ds} = 2.3$ V). Note that difference between measured and simulated $S_{21}$ is quite small in both cases. Nevertheless, Fig. 10 highlights that percentage error associated to $S_{21}$ under pinch-off condition is much higher than under a typical operating bias point. This is due to the small value of $S_{21}$ over the full frequency range, especially at low frequency, under such bias condition.
In this paper a comparison of two neural models of scaled microwave GaAs FETs is given. A comprehensive analysis is done, resulting in the following conclusions.

Both models provide a single bias-dependent model for all three modelled devices that can be further used as the circuit simulator library elements. This allows efficient simulation of S-parameters with arbitrary steps of bias voltages, which can be very important for building subsequently a large signal model. Further, unlike the standard modelling approaches that require the measured S-parameters for each considered bias point, both presented models require the measured S-parameters only for certain number of bias points to be used for the model development.

Since the black-box model is trained using the measured data, all effects contributing to the device behavior are included in the model, therefore it exhibits high modelling accuracy. On the other hand, the hybrid model retains the advantages of the equivalent circuit based model (here the analytical one), like the connection with the physical device structure, which represents a useful feedback for a quick and reliable optimization of the device fabrication processing. The accuracy of the equivalent circuit is of fundamental importance for developing a successful hybrid model.

While the hybrid approach requires certain number of extractions of the ECPs and then training of the corresponding ANNs, the black-box model requires only training of the ANNs. Hence, as far as the efficiency of the model development procedure is concerned, the black-box approach is simpler.

Both modelling approaches exhibit advantages and disadvantages. The best solution is to choose the approach depending on the specific case under study.

VI. CONCLUSION

REFERENCES


