Fuzzy-Neural Approach for Efficient Microwave Transistor Modeling

L. JI and M.C.E. YAGOUB
School of Information Technology and Engineering
University of Ottawa
800 King Edward Ave, Ottawa, Ontario, K1N 6N5
CANADA

Abstract: - The design of microwave systems requires accurate modeling of the behavior of active components such as MOSFETs, MODFETs, MESFETs, and HEMTs. This aspect is crucial when the design process leads to massive and highly repetitive computational tasks during simulation, optimization and statistical analysis. In this paper, we present a robust approach that combines the capabilities of neural networks and fuzzy theory to automatically predict the most reliable equivalent circuit model of microwave field effect transistors.

Key-Words: - Circuit Topology, FET Modeling, Fuzzy Neural Networks, Parameter Extraction.

1 Introduction

The rapid growth of today's microwave systems requires a continuous upgrading of existing models at the component level, taking into account emerging technologies and new applications. Widely used in the microwave area, field effect transistor (FET) gained a particular attention during the last decades. The most efficient approach is to model it as an equivalent electrical circuit which element values can be determined either by direct extraction [1] or by optimization-based extraction [2]. Fast and simple to implement, direct-extraction techniques provide adequate values for the more dominant circuit model elements but they cannot determine all the extrinsic elements uniquely [3]. On the other hand, optimization-based extraction techniques are more accurate but computationally intensive and sensitive to the choice of starting values [4]. Though several optimization-based extraction methods that are insensitive to starting values have recently been proposed, it is still difficult to determine all the model elements with a high degree of certainty. This is especially true for the extrinsic elements due to the small influence they have on the measured data and thus on the optimization process, making traditional optimizers to be numerically ill conditioned.

Furthermore, in order to make them attractive to non-experienced users, such techniques often assume a *prior* transistor circuit topology [5]. Considering the very large panel of FETs available on the market based on the fabrication process and final usage, this topology, referred in this paper as the *standard topology* (Fig. 1), cannot be suitable, especially when massive and highly repetitive computational tasks are required during design.

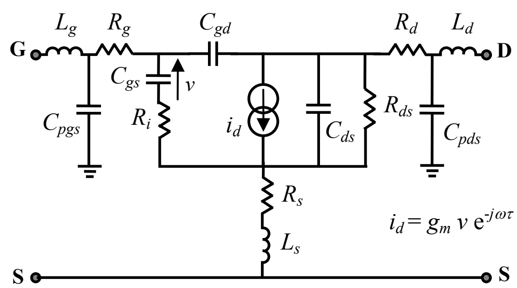

Fig. 1. FET standard topology.

Therefore, the aim of this paper is to present a robust approach that combines the capabilities of neural networks and fuzzy theory to automatically predict the most reliable circuit model of microwave FETs. Since the equivalent circuit of a FET is often specific to a given type of transistor and it is puzzling to decide which one is most suitable to be used in one specific application, we created a library of the most often used topologies, displayed in Fig. 2 to 5 [6]-[9].

Neural modeling is one of the most recent trends in microwave CAD. Fast, accurate and reliable neural network models can be trained from measured or simulated data. Once developed, these neural models can be used in place of CPU-intensive models to speed up design [10].

Fuzzy neural networks have been recently recognized as a useful vehicle for efficient modeling and design [11]. By combining the Fuzzy c-means method (FCM) [12] and the neural representation of a transistor behavior [13], the small-signal equivalent circuit parameters are efficiently evaluated through a fuzzy-neural network, based on an optimum selection of the more appropriate circuit topology of the active device.

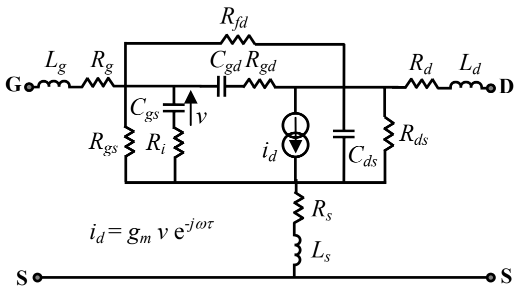

Fig. 2. Circuit topology # 1 as reported in [6].

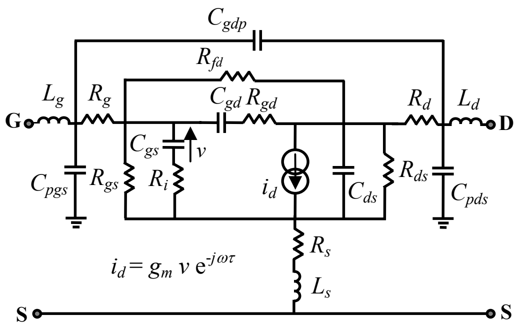

Fig. 3. Circuit topology # 2 as reported in [7].

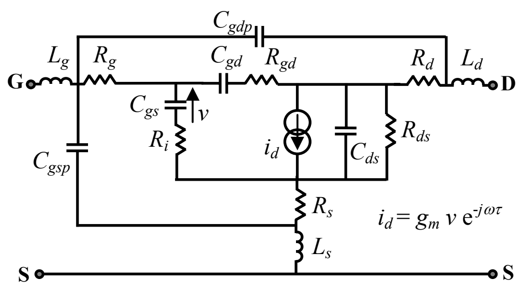

Fig. 4. Circuit topology # 3 as reported in [8].

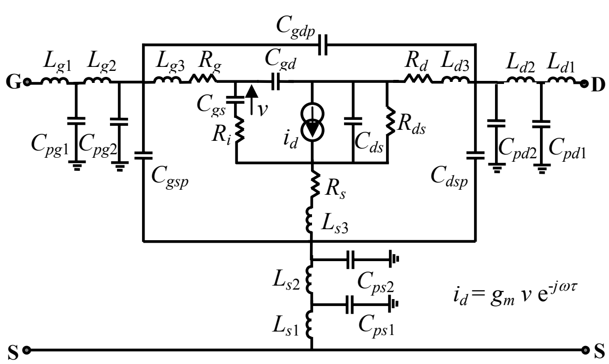

Fig. 5. Circuit topology # 4 as reported in [9].

2 Fuzzy-Neural Approach

The first step of the proposed method is a direct parameter extraction of the standard FET topology using the classical method described in [1]. The S-parameters ($S_{ij}^{\mathbf{s}}$, $i, j = 1, 2$) of the standard topology are then compared to the measured input parameters (denoted as $S_{ij}^{\mathbf{m}}$, $i, j = 1, 2$), as shown in Fig. 6. If the achieved accuracy is not acceptable, a new circuit topology should be selecting from the FET library.

FCM is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade [14]. Clustering in N unlabeled data X = $\{x_i,\ i = 1, \ldots, N\}$ is the assignment of c number of partition labels to the vectors in X. The problem of fuzzy clustering is to find the optimum matrix U = $[U_{ij} \in [0, 1],\ i = 1, \ldots, c;\ j = 1, \ldots, N]$ which minimize the function [14]

$$J_h(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{c} (U_{ik})^h \left\| x_k - v_i \right\|^2 \tag{1}$$

where h is an exponent that controls the degree of fuzziness, u_{ik} describes the belongness of x_i to cluster k,

$$u_{ik} = \left(\sum_{j=1}^{c} \left[\frac{\left\| x_k - v_i \right\|}{\left\| x_k - v_j \right\|} \right]^{\frac{2}{h-1}} \right)^{-1} \tag{2}$$

and v_i is the centroid of ith cluster,

$$v_i = \frac{\sum_{k=1}^{N} (u_{ik})^h\, x_k}{\sum_{k=1}^{N} (u_{ik})^h} \tag{3}$$

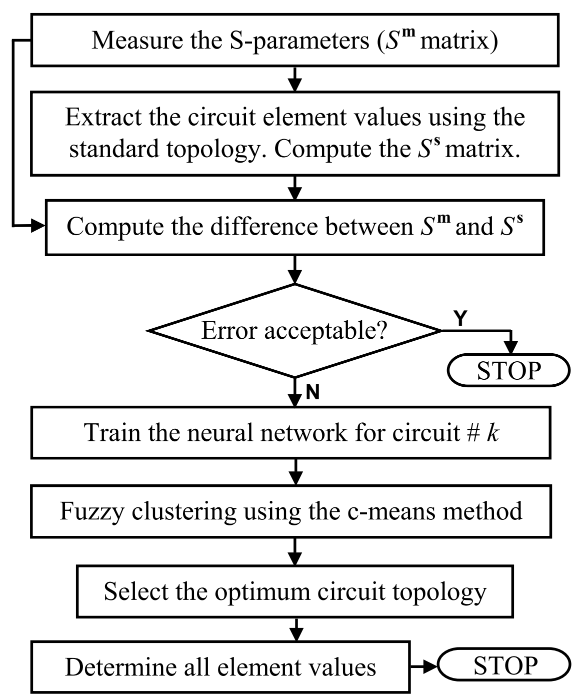

Fig. 6. Algorithm of the proposed method.

In this work, FCM would be an efficient tool to identify the optimum topology based on the following approach: For any circuit # k ($k = 1, \ldots, 4$), the related $S^{\mathbf{k}}$ matrix would be compared to the input $S^{\mathbf{m}}$ matrix and each element of the two resulting 2x2 error matrices $E^{\mathbf{k},\mathbf{Re}}$ and $E^{\mathbf{k},\mathbf{Im}}$,

$$E_{ij}^{\mathbf{k},\mathbf{Re}} = \mathrm{Re}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{m}}\right) \qquad i, j = 1, 2 \tag{4}$$

$$E_{ij}^{\mathbf{k},\mathbf{Im}} = \mathrm{Im}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{m}}\right) \qquad i, j = 1, 2 \tag{5}$$

would receive a score scaled from 1 to 10 depending on its value. Thus, topology #k with smallest $E^{\mathbf{k},\mathbf{m}}$,

$$E^{\mathbf{k},\mathbf{m}} = \sum_{i=1}^{2}\sum_{j=1}^{2}\left\{\left[\mathrm{Re}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{m}}\right)\right]^2 + \left[\mathrm{Im}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{m}}\right)\right]^2\right\} \tag{6}$$

i.e., smallest score, would be selected as the most adequate circuit. In the above equation, Re(*) and Im(*) denote the real part and the imaginary part respectively. However, this approach is not practical. In fact, since there is *no prior* knowledge on the input parameters, it is impossible to compute numerically (6). Let $\{\Omega^{\mathbf{s}}\}$ be the set of $P_{\mathbf{s}}$ elements $\Omega_p^{\mathbf{s}}$ ($p = 1, \ldots, P_{\mathbf{s}}$) in the standard topology. A symbolic code was developed using [15] to analytically derive the following nonlinear functions

$$S_{ij}^{\mathbf{k}} = f_{ij}^{\mathbf{k}}\left(S_{ij}^{\mathbf{s}}, \left\{\Omega^{\mathbf{k}}\right\}\right) \quad i, j = 1, 2 \quad k = 1, \ldots, 4 \tag{7}$$

where $\{\Omega^{\mathbf{k}}\}$ is the set of the $P_{\mathbf{k}}$ elements added in circuit #k in comparison with the standard topology,

$$\left\{\Omega^1\right\} = \left\{R_{fd}, R_{gd}, R_{gs}\right\}\Big|_{C_{pgs}=C_{pds}=0}$$

$$\left\{\Omega^2\right\} = \left\{R_{fd}, R_{gd}, R_{gs}, C_{gdp}\right\}$$

$$\left\{\Omega^3\right\} = \left\{R_{gd}, C_{gdp}, C_{gsp}\right\}\Big|_{C_{pgs}=C_{pds}=0}$$

$$\left\{\Omega^4\right\} = \left\{\begin{matrix} C_{gdp}, C_{gsp}, C_{dsp}, C_{pg1}, C_{pg2}, \\ C_{ps1}, C_{ps2}, C_{pd1}, C_{pd2}, L_{g1}, \\ L_{g2}, L_{s1}, L_{s2}, L_{d1}, L_{d2} \end{matrix}\right\}\Bigg|_{C_{pgs}=C_{pds}=0}$$

in order to evaluate the alternative fuzzy criteria

$$E^{\mathbf{k},\mathbf{s}} = \sum_{i=1}^{2}\sum_{j=1}^{2}\left\{\left[\mathrm{Re}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{s}}\right)\right]^2 + \left[\mathrm{Im}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{s}}\right)\right]^2\right\} \tag{8}$$

The above relations depend *only* on the values of the $P_{\mathbf{k}}$ elements of set $\{\Omega^{\mathbf{k}}\}$. However, relations (8) are strongly interdependent, highly nonlinear, multi-dimensional, and require a huge combination of values to be accurately evaluated. Therefore, we used neural networks to learn these quantities. A neural network (NN) is a model that has the ability to learn and generalize arbitrary continuous multi-dimensional nonlinear input-output relationships. Let $\mathbf{x}$ be an n-vector $\{x_i, i = 1, \ldots, n\}$ containing the inputs and $\mathbf{y}$ be an m-vector $\{y_k, k = 1, \ldots, m\}$ containing the outputs from the output neurons. The original problem can be expressed as $\mathbf{y} = \boldsymbol{f}(\mathbf{x})$, while the neural network model for the problem is

$$\mathbf{y}_{NN} = \tilde{\mathbf{y}}(\mathbf{x}, \mathbf{w}) , \tag{9}$$

where $\mathbf{w}$ is a N_w-vector $\{w_i, i = 1, \ldots, N_w\}$ containing all the weight parameters representing the connections in the NN. The definition of $\mathbf{w}$ and the way in which $\mathbf{y}_{NN}$ is computed from $\mathbf{x}$ and $\mathbf{w}$ determines the structure of the NN. The most commonly used NN configuration is the Multi Layer Perceptrons (MLP) where the neurons are grouped into layers as shown in Fig. 7.

The layer L_1 is the input layer. The layers L_2 to L_{L-1} are called hidden layers, while the last layer L_L, the output layer, contains the response to be modeled. The various layers are placed end to end with neuron connections between them. For such a neural network, the function given by (9) is calculated on the basis of the layer of entry while using [10]

$$z_i^1 = x_i , \quad i = 1, \ldots, N_1, \quad n = N_1 \tag{10}$$

z_i^1 is the output of the i^{th} neuron of the input layer, and while proceeding layer by layer, the output at the end of layer L_l is given by

$$z_j^l = \sigma\left(\sum_{k=0}^{N_{l-1}} w_{jk}^l \; z_k^{l-1} + w_{j0}^l\right), \tag{11}$$

where $j = 1, \ldots, N_l$, and $l = 1, \ldots, L$, to reach the output layer that gives

$$y_k = z_k^L , \quad k = 1, \ldots, N_L, \quad m = N_L \tag{12}$$

In these relations, N_l is the number of neurons in the layer L_l, w_{jk} represents the weight of the connection between the k^{th} neuron of the layer L_{l-1} and the j^{th} neuron of the layer L_l.

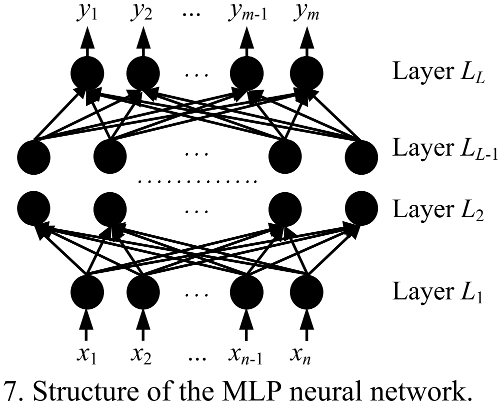

Fig. 7. Structure of the MLP neural network.

In (11), the function σ is known as the activation function of the neuron. By allocating values to the standard $S^{\mathbf{s}}$ parameters and varying the value of each element $\Omega_p^{\mathbf{k}}$ (p = 1, …, $P_{\mathbf{k}}$) of set $\{\Omega^{\mathbf{k}}\}$, we utilized (7) to compute the $S^{\mathbf{k}}$ parameters and therefore, the difference $\{S^{\mathbf{k}} - S^{\mathbf{s}}\}$. The resulting data in the form of

$$Tr^{\mathbf{k}} = \left[\underbrace{\mathrm{Re}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{s}}\right), \mathrm{Im}\left(S_{ij}^{\mathbf{k}} - S_{ij}^{\mathbf{s}}\right)}_{8\,\text{inputs}\ (i,\,j=1,2)}, \underbrace{\Omega_1^{\mathbf{k}}, \ldots, \Omega_{p_{\mathbf{k}}}^{\mathbf{k}}}_{p_{\mathbf{k}}\ \text{outputs}} \right] \tag{13}$$

was submitted to a three-layer (MLP3) neural network structure for training using the *Neuromodeler* tool [16] as shown in Fig. 8.

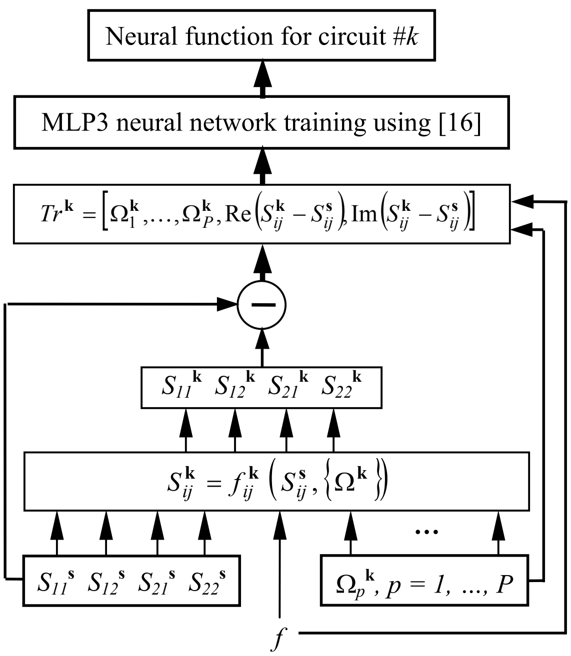

Fig. 8. Neural network model for circuit #*k.*

The input layer has 9 neurons (the 4 real and 4 imaginary parts in (8) and the operating frequency *f*) while the output layer contains $P_{\mathbf{k}}$ neurons. The hidden layer is composed of 22 to 45 neurons depending on the circuit data file under training.

It has to be noted that once the inputs and outputs are identified, three sets of data namely, the training data, the validation data, and the test data, need to be generated for the neural network development. Training data is used to guide the training process, i.e., to update the neural network weights during training. Validation data is used to monitor the quality of the neural model during training. Test data is used to examine the final quality of the developed model.

Suppose the range of input parameters over which the neural model would be used is $[\boldsymbol{x}_{min}, \boldsymbol{x}_{max}]$. Therefore, validation data, test data, and training data should be generated in the same range as well, selecting a sampling strategy and an adequate step size. Grid, star, central-composite, or random distributions are possible.

Prior to further discussion, two points had to be considered in this work. First, the input parameter space is of high-dimension and the step sizes should be small enough to assure good convergence. This will lead to a too large number of combinations of input parameters. Second, even after selecting the optimum topology, the values of the elements of set $\{\Omega^{\mathbf{s}}\}$ obtained after the first round of extraction, i.e., using the standard topology need to be tuned in the final circuit along with the neural outputs, i.e., the elements of set $\{\Omega^{\mathbf{k}}\}$. An optimization loop is then essential. Therefore, instead of generating large data files required for a classical neural development, we used the values of the following vector

$$\Omega = \left[\Omega_1^{\mathbf{k}}, \ldots, \Omega_{P_{\mathbf{k}}}^{\mathbf{k}}, \Omega_1^{\mathbf{s}}, \ldots, \Omega_{P_{\mathbf{s}}}^{\mathbf{s}}\right] \tag{14}$$

as starting vector for the optimization loop. This procedure will assure better and faster convergence.

3 Validation

The device to be characterized is the one reported in [17] using topology # 3. Since in this paper all circuit element values are given as well as the final error between measured and simulated *S*-parameters, a reliable comparison can be performed for a full validation. In fact, by comparing the S-parameters (Fig. 9) and the element values (Table I) in [17] with those obtained in 2.3 seconds using our technique, circuit # 3 achieved a quite close agreement as expected, with a smaller final error.

The second device to be characterized is FET EPA018A. After 2.1 seconds, our method showed that circuit #2 is the most appropriate (Fig. 10) with an acceptable final error of 1.8%, smaller than the user specifications, i.e., 2%.

4 Conclusion

In this paper, a combined fuzzy-neural tool has been used for determining the optimum small-signal FET equivalent circuit topology. The method has been proven to be fast and accurate and can be applied to other RF/microwave active devices such as HBTs.

Acknowledgement

This work is supported in part by Natural Science and Engineering Research Council of Canada.

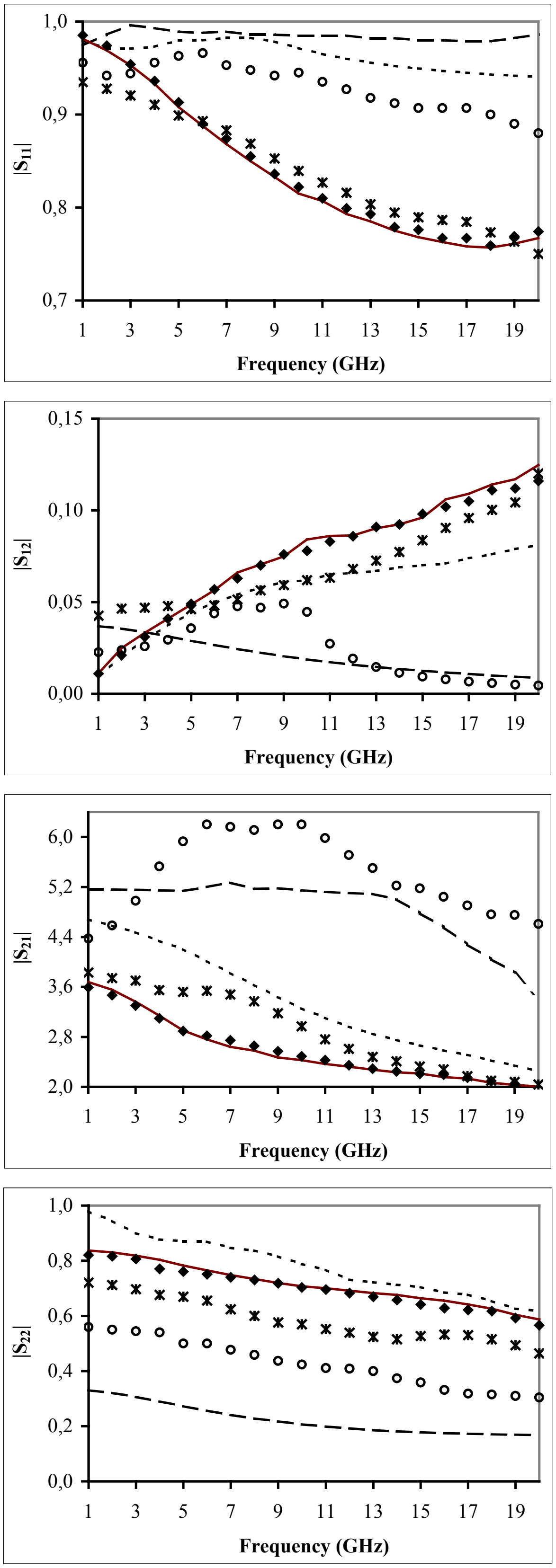

Fig. 9. Comparison of measured S-parameters (♦) with those extracted using different topologies:
---- : standard topology, —— : topology # 1,
o : topology # 2, —— : topology # 3, * : topology # 4.

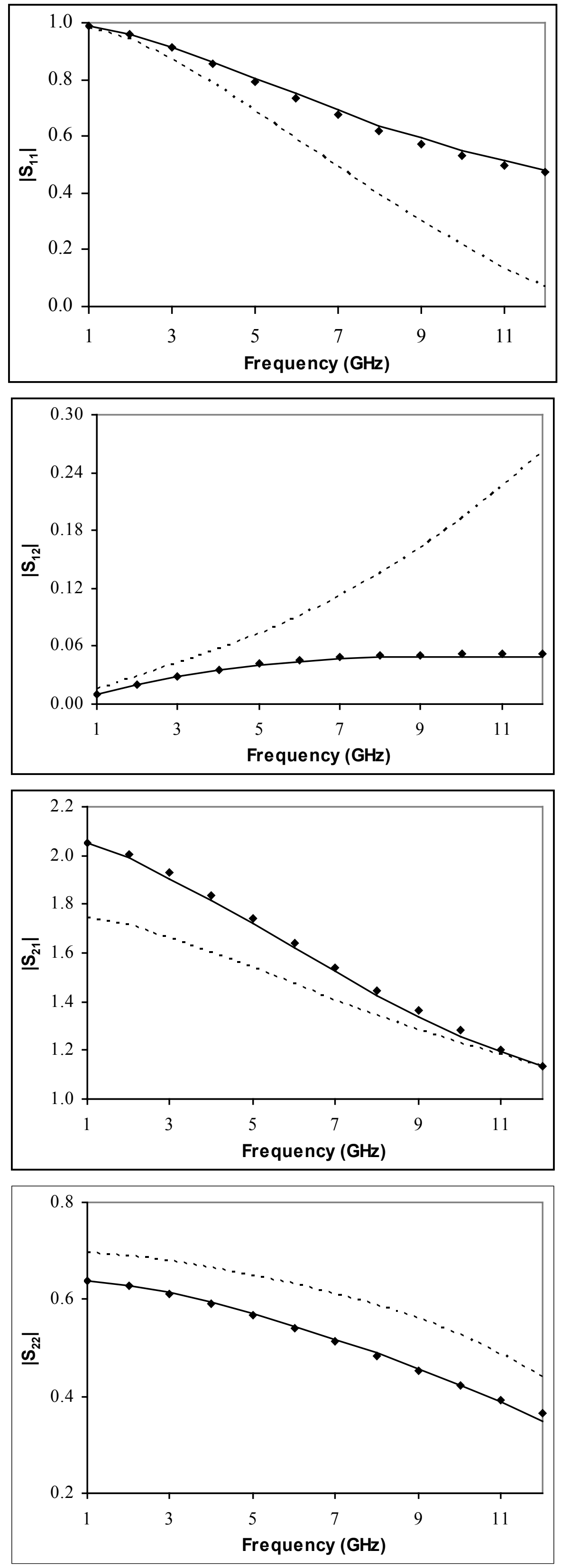

Fig. 10. Comparison of Measured S-Parameters (♦) with those extracted using different topologies:
---- : standard topology,
—— : topology # 2.

Table I. Comparison between the parameters reported in [17] and our results for example 1.

	Circuit # 3	Our Values
C_{gs} (pF)	0.277	0.215
C_{gd} (pF)	0.0207	0.0211
C_{ds} (pF)	0.0993	0.101
g_m (mS)	26.9	27.3
τ (ps)	1.22	1.25
R_i (Ω)	15.3	15.1
R_{gd} (Ω)	43.8	43.6
R_{ds} (Ω)	215	218
R_g (Ω)	8.9	9.1
R_s (Ω)	7.5	7.3
R_d (Ω)	13.6	13.2
L_s (nH)	0.437	0.441
L_d (nH)	0.452	0.447
L_g (nH)	0.254	0.258
C_{gsp} (pF)	0.0409	0.0397
C_{gdp} (pF)	0.001	0.001
Error (%)	8.4	2.9

References:

[1] J.M. Golio, *Microwave MESFETs and HEMTs*, Boston: Artech House, 1991.

[2] L. Fujiang, G. Kompa, "FET model parameter extraction based on optimization with multiplane data-fitting and bidirectional search-a new concept," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1114-1121.

[3] C. Van Niekerk, P. Meyer, D.M.M.-P. Schreurs, P.B. Winson, "A robust integrated multibias parameter-extraction method for MESFET and HEMT models," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 777-786.

[4] C. Van Niekerk, J.A. du Preez, D.M.M.-P. Schreurs, "A new hybrid multibias analytical / decomposition-based FET parameter extraction algorithm with intelligent bias point selection," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 2003, pp. 893-902.

[5] G. Dambrine, A. Cappy, F. Heliodore, E. Playez, "A new method for determining the FET small-signal equivalent circuit," *IEEE Trans. Microwave Theory Tech.*, Vol. 36, 1988, pp. 1151-1159.

[6] M. Berroth, R. Bosch, "High frequency equivalent circuit of GaAs depletion and enhancement FETs for large signal modelling," *Workshop on Measurement Techniques for Microwave Device Characterization and Modelling*, 1990, pp. 122-127.

[7] R. Menozzi, A. Piazzi, F. Contini, "Small-signal modeling for microwave FET linear circuits based on a genetic algorithm," *IEEE Trans. Circuits and Systems*, Vol. 43, 1996, pp. 839-847.

[8] M. Fernandez-Barciela, P.J. Tasker, Y. Campos-Roca, M. Demmler, H. Massler, E. Sanchez, M.C. Curras-Francos, M. Schlechtweg, "A simplified broad-band large-signal nonquasi-static table-based FET model," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 395-405.

[9] ADSv.2003, *Agilent Technologies*, Palo Alto, CA, USA

[10] Q.J. Zhang, and K.C. Gupta, *Neural Networks for RF and Microwave Design*, Norwood: Artech House, 2000.

[11] Y. Hu, K. Ashenayi, R. Veltri, G. O'Dowd, G. Miller, R. Hurst, R. Bonner, "Comparison of neural network and fuzzy c-means methods in bladder cancer cell classification," *IEEE Int. Conf. on Neural Networks*, 1994, pp. 3461-3466.

[12] S. Miyamoto, "An overview and new methods in fuzzy clustering," *Int. Conf. on Knowledge-Based Intelligent Electronic Systems*, 1998, pp. 33-40.

[13] B. Karlik, H. Torpi, M. Alci, "A fuzzy-neural approach for the characterisation of the active microwave devices," *Int. Conf. Microwave and Telecommunication Technology*, 2002, pp. 114-117.

[14] C. Fager, L.J.P. Linner, J.C. Pedro, "Optimal parameter extraction and uncertainty estimation in intrinsic FET small-signal models," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, 2002, pp. 2797-2803.

[15] B.W. Char, K.O. Geddes, G.H. Gonnet, B.L. Leong, M.B. Monagan, S.M. Watt, *Maple V reference manual*, New-York: Springer, 1991.

[16] *NeuroModeler Version 1.02*, Prof. Q.J. Zhang, Carleton University, Ottawa, ON, Canada.

[17] M.K. Ahmed, S.M.M. Ibrahem, "Small signal GaAs MESFET model parameters extracted from measured S-parameters," *National Radio Science Conference*, Cairo, Egypt, 1996, pp. 507-515.