PRIOR KNOWLEDGE INPUT METHOD IN DEVICE MODELING

Serdar Hekimhan Serdar Menekay N. Serap Sengör

e-mail: shekimhan@hotmail.com, smenekay@mynet.com, neslihan@ehb.itu.edu.tr

Istanbul Technical University, Faculty of Electrical & Electronics Engineering, 80626, Maslak, Istanbul, Turkey

ABSTRACT

In this paper, Multi-Layer Perceptron and wavelet-based neural networks are utilized in Prior Knowledge Input method to model BJT. The benefit of Prior Knowledge Input method over plain usage of Artificial Neural Networks in modelling is investigated by comparing models obtained with or without PKI. Training and test data used during simulations are obtained by HP 4155 parameter analyser.

I. INTRODUCTION

Electronic devices' response can be obtained by solving physical equations using simulators as Electromagnetic (EM) simulators at RF/microwave frequencies or measurement. As these techniques are time consuming in obtaining responses, Artificial Neural Networks (ANN), which give less accurate results than EM simulators but more accurate results than conventional methods as empirical models, look-up tables etc., is introduced to eliminate drawback of time consumption in EM simulators and non-accuracy of conventional techniques during CAD based design [1]. This approach is becoming more common as applications at RF/microwave frequencies become more needed. Besides its comparable fastness and accuracy, ANN has the advantage of giving black-box models in case of insufficient knowledge to obtain a model of a device.

The aim of this work is to investigate PKI method, especially with different ANN structures and observe its advantages over using plain ANN structures in modelling. In this work, as an example transistor is modelled because transistors, especially BJT's, are widely used in discrete circuits as well as in IC design, in both analog and digital applications. BJT has different responses at various frequencies. In order not to solve different equations at different frequencies, ANN can be used to obtain its response at a broad spectrum. However in this work as training and test sets are not sufficient, only model at DC working conditions are obtained. This is a drawback of ANN, as models depend on training set dramatically.

However, in literature, various methods based on neural networks such as space mapping [2], knowledge-based [3], prior knowledge input [4,5] are proposed to overcome this inconvenience and to prevent long training time. On account of its simplicity Prior Knowledge Input (PKI) is considered in this work to obtain a model of BJT to prevent long training.

In section II, Multi-Layer Perceptron (MLP) [6] and Wavelet-based neural networks [7] used in obtaining models of devices and circuits especially used at RF/microwave frequencies are presented. In section III PKI method is introduced and in section IV simulation results obtained are given.

II. MLP and Wavelet-based Neural Network

In most applications of ANN in modelling, a feed forward neural network is used to get a non-linear input-output mapping related with a device or circuit [1].
In this paper two kinds of feed forward neural networks, MLP and wavelet-based, will be presented. Besides using these ANN structures in PKI method, for comparison MLP based modelling will also be given.

2.1 Multi-layer Perceptrons

MLP as shown in Fig.1 has three layers input, hidden and output, respectively. Hidden layers, which can be more than one, are formed using neurons having continuously differentiable non-linear activation functions, while output neurons have linear activation functions for MLP used in modelling. Input layer has no function than fan out the data. Input data move ahead through the layers and at output layer, output of the ANN is obtained. These outputs, y, are compared to desired values, y_d, and weights are changed by means of back propagation algorithm minimizing error function depending on error, e, between desired and obtained outputs. Error, instantaneous error of k-th training data and error function are given below by Eq. (1), (2) and (3), respectively.

$$e = y_d - y \tag{1}$$

$$e^{(k)} = \frac{1}{2}\sum_j e_j^{(k)^2} \tag{2}$$

$$e_{ort} = \frac{1}{p}\sum_{k=1}^{p} e^{(k)} \tag{3}$$

Changing weights is carried on mostly until average error of training set, is in a predetermined range. Stopping criteria given in Eq. (4) is taken for this work where change in average error is considered rather than average error itself.

$$\left|\Delta e_{ort}\right| < e \tag{4}$$

As backpropagation algorithm is well-known [6] it will not be mentioned further here.

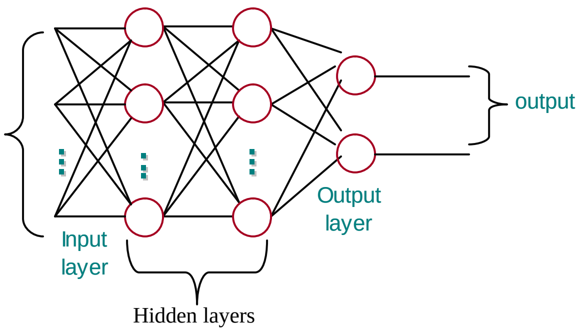

Figure 1. Multilayer Perceptron structure.

2.2 Wavelet Based Network

Another feed-forward structure used as universal approximator in different applications is wavelet-based neural network. MLP's hidden layers can be more than one but wavelet-based types have only one hidden layer as shown in Fig. 2. Wavelet-based neural network formed by incorporating wavelet theory into neural network has been a new type [7]. Difference between MLP and wavelet-based is that activation functions of hidden layer of neural network are wavelet functions. Generally used inverse Mexican hat function given in Eq. (5), (6) is selected in this work.

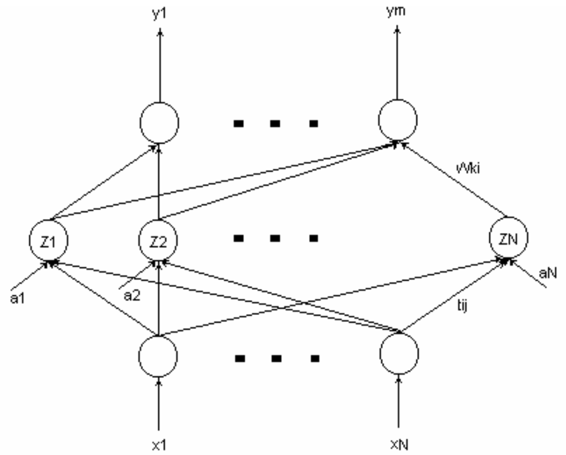

Figure 2. Wavelet-based neural network structure.

$$Z_i = \Psi\left(\frac{x - t_i}{a}\right) = s\,(g_i) = (g_i^2 - n)\exp\left(-\frac{g_i^2}{2}\right) \tag{5}$$

$$g_i = \left\|\frac{x - t_i}{a}\right\| = \sqrt{\sum_{j=1}^{n}\left(\frac{x - t_i}{a}\right)^2} \tag{6}$$

Outputs are obtained by Eq. (7), where n and m denote the number of outputs and hidden neurons, respectively.

$$y_i = \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} z_j \tag{7}$$

$$W_{ji}^{(k+1)} = w_{ji}^{(k)} - h\frac{\partial E^{(k)}}{\partial w_{ji}^{(k)}} \tag{8}$$

$$a_i^{(k+1)} = a_i^{(k)} - h\frac{\partial E^{(k)}}{\partial a_i^{(k)}} \tag{9}$$

$$t_{ij}^{(k+1)} = t_{ij}^{(k)} - h\frac{\partial E^{(k)}}{\partial t_{ij}^{(k)}} \tag{10}$$

Outputs are compared to desired outputs same as in MLP, and parameters w_{ij}, t_{ij} and a_{ij} are changed to minimize error function given in Eq. (2) using steepest descent method. As back-propagation algorithm depends on the same optimisation method updating of parameters by Eqs. 8-10 are very similar to back-propagation but this time as different type of parameters are of concern at each layer there is no need to back-propagate error. ? in these equations is the learning rate. When Eq. (4) is satisfied, training is stopped. In order to understand whether network is well trained, it has to be tested with test data.

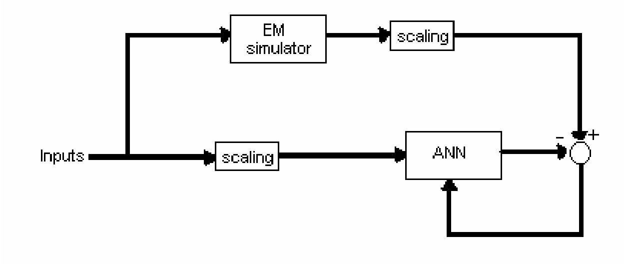

Figure 3.General neural network structure during training phase for model generating.

How ANN structures are trained for modelling purpose is shown in Fig.3. Scaling unit is needed to carry the circuit/device parameters and responses into a range meaningful for ANN structure considered. Once training is over, the weights and bias terms of trained ANN is saved and ANN model is used for design purposes. It has to be pointed out that even training phase is long, using ANN model it takes very short time as there is no complex calculations.

III.PRIOR KNOWLEDGE INPUT (PKI) METHOD

For good training of neural networks large amount of data are needed. But, on the other hand, training time increases with the number of data. Furthermore, obtaining large set of data is not an easy process as it depends either on measurement or time taking difficult computations. So that if any information exists about component, circuit etc., which is to be modelled, training time can be reduced with addition of knowledge into modelling method. PKI is such a method that, it not only reduces training time but also produces more accurate results than method of modelling based on plain ANN structures.

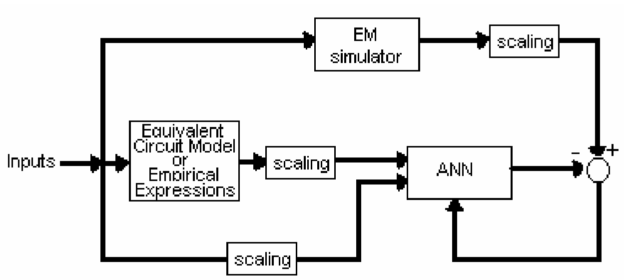

Figure 4. PKI Neural Network structure

PKI method was first presented by Gupta. In this method inputs of ANN are no longer only physical/process parameters related with the device to be modelled but also outputs of equivalent circuit model/empirical equations as shown in Fig. 4. This is the main difference of the PKI compared to ANN method without PKI. This method was used for modelling finline [4] empirical expressions are used as prior knowledge, and for microwave components [5] where an ANN model trained beforehand is used as knowledge.

The difference between using ANN with PKI and without PKI during training can be seen from Fig. (3) and (4). In case of PKI method inputs to ANN is enlarged, but as will be seen in simulation results this has no disadvantage on training phase, indeed training phase is dramatically shortened. Once training is over model obtained by PKI is composition of two sub-systems, equivalent circuit/empirical relations and ANN.

IV. SIMULATION RESULTS

It is expected that ANN is more accurate than emprical relations and faster than fine model. To observe this MLP based ANN structure explained in section 2.1 is trained and model of BJT is obtained. Also to benefit the knowledge about BJT given by empricial relations PKI method is utilized. BJT Ebers-Moll 3 model is considered as emprical relations to obtain collector and base current and two different models are obtained by PKI method where two different ANN structures, namely MLP and wavelet-based, are used. Thus on total three models for BJT are obtained and these are compared with experimental data and model obtained by emprical relations.

Desired output data is obtained by HP 4155 parameter analyser. After data is obtained, they must be scaled. Scaling plays very important role for the ANN since if data is not seperately distributed in the region of interest, formed according to activation functions used covergence may not occur. So that in this study, firstly logarithmic scaling and then linear scaling, given by Eq. (11), (12) respectively, are used. Scaling limits are -0.9 and 0.9 for MLP and -1 and 0.5 for wavelet-based type. As mentioned above these limit values depend on activation functions used so, since in MLP with and without PKI, hyperbolic tangent function is used limit values are –1 and 1. The range of inverse Mexican hat function is on the interval of [-1, 0.5] the limit values are choosen appropriately.

$$\hat{X} = \ln(X) \tag{11}$$

$$\hat{X} = \hat{X}_{\min} + \frac{X - X_{\min}}{X_{\max} - X_{\min}}\left(\hat{X}_{\max} - \hat{X}_{\min}\right) \tag{12}$$

where,

X : value taken from data set

$\hat{X}$: scaled value

X_{min} : minimum of the data set

$\hat{X}_{min}$:minimum scaled value

X_{max} :maximum of the data set

$\hat{X}_{max}$: maximum scaled value

One of inputs of neural networks scaled is base-emitter voltage, and other two inputs scaled are collector and scaled base currents obtained by Ebers-Moll 3 model given in Eqs. (13)-(14). Due to PKI method ANN structure is in a way restricted with equations (13)-(14).

$$I_c = I_s(0)\left[\exp\left(\frac{V_{BE}}{V_T}\right) - 1\right] \qquad (13)$$

$$I_b = \frac{I_s(0)}{b_{FM}}\left[\exp\left(\frac{V_{BE}}{V_T}\right) - 1\right] + C_2 I_s(0)\left[\exp\left(\frac{V_{BE}}{n_{EL}.V_T}\right) - 1\right] \qquad (14)$$

where,

$I_s = 69{,}144.10^{-14} A$

$V_T = 25{,}875.10^{-3} V$

$b_{FM} = 506$

$C_2 = 74{,}165$

$n_{EL} = 2.94$

The simulations are done using codes written in MATLAB® as m-files. Neural network toolbox of MATLAB® is not used since a composite system special for modeling purpose is formed. This same system can be used in obtaining models for different circuits/devices once training/test sets and for PKI method empricial equations/equivalent circuit are given.

In order to have meaningful comparison all ANN structures have same nimber of neurons in their hidden layer, namely 10 neurons. It has been tried with different number of neurons and observed that results are getting worse once the number of hidden layer neurons is over 18. This is due to over parameterization. Inputs of ANN in PKI method are more than without PKI method, they are 3 and 1, respectively. Outputs are 2, corresponding to base and collector currents. Learning rate ? is choosen to be 0.5 as there had not been much change trying in the range of [0.3, 0.7].

The comparison of base currents obtained from ANN with and without PKI with measurement values are shown in Fig.(5),(6),(7). The comparison of training and test phases are given in Table (1) and (2), respectively. From Table (1), it is seen that there is a dramatic change in number iterations during training on the behalf of PKI method. It also has to be pointed out that one iteration takes less time in PKI, so training time is shortened.

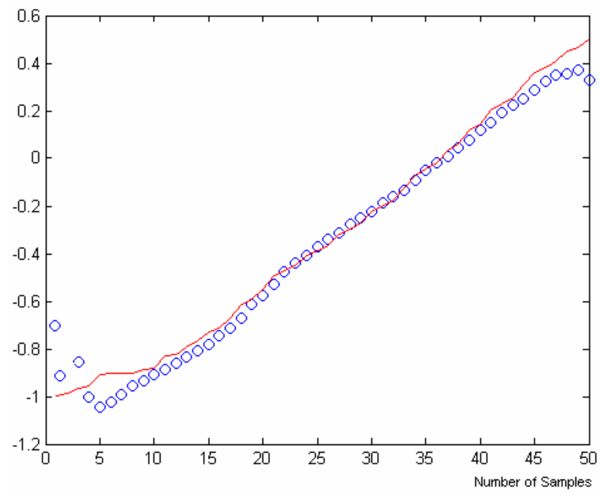

Figure 5. Ib values, with MLP (-desired -obtained results)

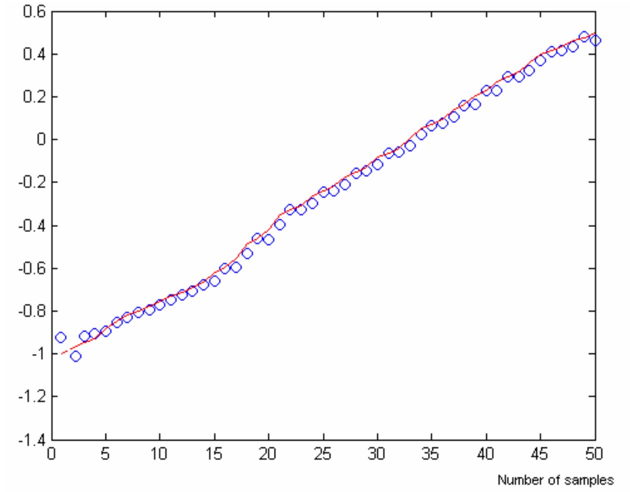

Figure 6. Ib values, with wavelet-based type (-desired -obtained results)

The comparison of results obtained from all three ANN based models with measurement and empirical equation is given in Fig. 8.

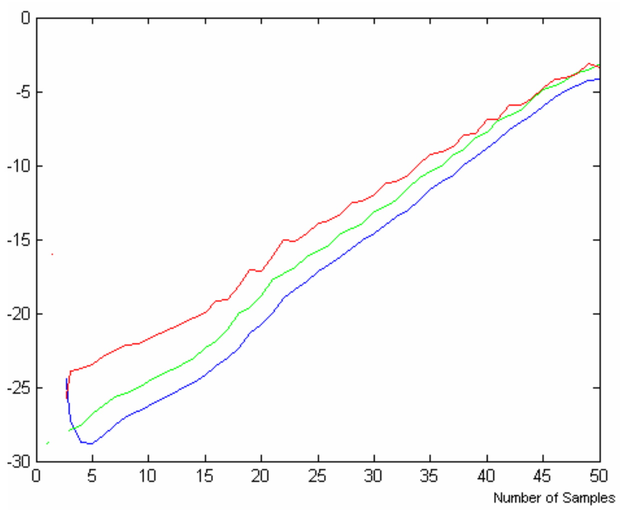

Figure 7. Logarithmic Ic values (-Wavelet-based type -MLP, --real results)

	Average error	Max. error	iteration
MLP without PKI	$1.84*10^{-2}$	0.0425	27
MLP with PKI	$1.7279*10^{-3}$	$3.019*10^{-2}$	8
Wavelet-Based with PKI	$5.312*10^{-4}$	$3.101*10^{-3}$	8

Table 1. Training set values.

	Average error	Max. error
MLP without PKI	$2.9*10^{-2}$	0.01874
MLP with PKI	$1.104*10^{-3}$	$2.1061*10^{-2}$
Wavelet-based with PKI	$7.805*10^{-4}$	$4.525*10^{-3}$

Table 2. Test set values.

V. CONCLUSION

As simulation results point out in figures (5) to (7), PKI method gives better results than using empirical equations and ANN without PKI. This can be further observed from tables where also it is observed that generalisation of PKI method is superior over ANN without PKI. The most important advantage of PKI method is short training time. Even though modelling BJT is not a big challenge, the proposed advantages of PKI method in literature are recognised. In this work besides using MLP in PKI method, wavelet-based network is also utilized and it is observed that there is no big difference between two ANN structures. Moreover, for this application it is observed that MLP can be preferred over wavelet-based network in PKI.

REFERENCES

1. Zhang Q.J., K. C. Gupta, Neural Networks for RF and Microwave Design , Arctech House, 2000.
2. space mapping
3. Wang F., Zhang Q.J., Knowledge-Based Neural Models for Microwave Design, IEEE Trans. on Microwave Theory and Techniques, Vol. 45, No.12, December 1997.
4. Chao Li, Jun Xu, and Liangjin Xue, Knowledge-Based Artificial Neural Network Models for Finline, International Journal of Infrared and Millimeter Waves, Vol. 22, No. 2, 2001
5. Watson P.M., Gupta K.C., Mahajan R. L., Applications of Knowledge-Based Artificial Neural Network Modeling to Microwave Components, Int. J. RF and Microwave CAE, 9:254-260, 1999.
6. Haykin, S. Neural Networks: A Comprehensive Foundation, Prentice Hall, 1999.
7. Zhang Q., Using Wavelet Network in Nonparametric Estimation, IEEE Trans. On Neural Networks, Vol.8, No. 2, march 1997.