A Coherent Small/Large Signal FET model Based on Neuronal Architectures[1]

A. Mediavilla, A. Tazón, J.A. García, J.M. Zamanillo, T. Fernandez

Dpto. Ingeniería de Comunicaciones
ETSI TELECOMUNICACION, University of CANTABRIA
Avda. Los Castros, 39005 – Santander - SPAIN
Phone: +34-942-201490, Fax: +34-942-201488
Email: media@dicom.unican.es

ABSTRACT

A modular neural network architecture for accurate small/large signal microwave MESFET/HEMT modeling is presented. This is achieved by means of an original neural architecture having two main modules. A network captures the nonlinear dynamic Pulsed I/V characteristic of the device, which is mainly responsible of the large signal behavior, while the second network estimates the high order derivatives at the operation point, which are responsible of the IMD behaviour, by means of a neural network and then it locally reconstructs the current function by means of a third order Taylor series around that point. Finally, in order to have a maximum of coherence, the two networks are combined into a global model by means of a simple fuzzy controller. Computer simulations and experimental measurements validate this flexible modeling technique.

INTRODUCTION

It is well known that in a microwave FET device the predominant nonlinear element is the drain-to-source current I_{ds}, which depends on both the static control voltages (V_{dso}, V_{gso}) and the dynamic voltages superposed to the bias point (v_{ds}, v_{gs}). Thus, the instantaneous control signals would be the sum of both components, that is, $V_{ds} = V_{dso} + v_{ds}$ and $V_{gs} = V_{gso} + v_{gs}$. Furthermore, depending on the level of the dynamic voltages there are two clearly different regimes: large and small signal. To model the large-signal behavior, it is enough to accurately characterize the nonlinear dynamic pulsed *I/V* characteristic [1], but in a small-signal situation it is necessary to have a higher level of detail [2]. As it has been shown, the nth-order intermodulation output power varies as the square of the nth derivative of the *I/V* characteristic. This means that an accurate control if IMD distortion needs an accurate fit for the main nonlinear function as well as for its derivatives. In most of cases, the local behavior of I_{ds} is well described by:

$$I_{ds}^{SS)} = I_{dso} + G_m v_{gs} + G_{ds} v_{ds} + G_{m2} v_{gs}^2 + G_{md} v_{ds} v_{gs} + G_{d2} v_{ds}^2 + G_{m3} v_{gs}^3 + G_{m2d} v_{ds} v_{gs}^2 + G_{md2} v_{ds}^2 v_{gs} + G_{d3} v_{ds}^3 \quad (1)$$

where I_{dso} is the static drain current and (G_m, ..., G_{d3}) are coefficients related to the nth-order derivatives of the *I/V* characteristic with respect to the instantaneous voltages evaluated at the bias point. Therefore, our small-signal modeling problem consists of fitting a function (model) $g: \Re^2 \to \Re^{10}$, which approximates the nonlinear mapping from the input space of bias voltages $\boldsymbol{V}$ = (V_{dso}, V_{gso}) to the output space of coefficients of the Taylor expansion $g(\boldsymbol{V}) = (I_{dso}, G_m, G_{ds}, G_{m2}, G_{md}, G_{d2}, G_{m3}, G_{m2d}, G_{md2}, G_{d3})$. Once this model is available, the drain current will be reconstructed by using (1) the above truncated Taylor series expansion. Up to now, these two regimes are treated separately. In particular, we have recently proposed two different neural network structures to solve the both modeling problems: a smoothed piecewise-linear (SPWL) structure is used to model the large-signal behavior [3], and a generalized radial basis function (GRBF) network is used to estimate the function derivatives at the bias point, in order to characterize the small-signal behavior [4]. In this paper we combine in a consistent way these two networks into a single global model by means of fuzzy membership functions. In this way, the overall network provides a smooth transition between both regimes of behavior.

THE SMALL SIGNAL BEHAVIOUR

To obtain the small-signal mapping described above, we have applied a generalized radial basis function (GRBF) network [4], which is an extension of the well-known RBF network that relaxes the radial constraint of the Gaussian kernels thus allowing different variances for each dimension of the input space and leading to elliptic basis kernels. This network seems specially suited in this application because of the shape of the coefficients G_m ... G_{d3}: while the dependence with V_{dso} is quasi-linear, the dependence with V_{gso} suggests that they could be approximated by a combination of Gaussians. The output of the GRBF network is given by:

$$g(\boldsymbol{V}) = \sum_{i=1}^{I} g_i(\boldsymbol{V}) \quad (2)$$

[1] This work has been partially supported by CYCIT grant 1FD97-1863-C02-01 and Accion Integrada HF1999-63

where i indexes the GRBF units, $g_i(\boldsymbol{V}) = \lambda_i\, o_i(\boldsymbol{V})$, while $o_i(\boldsymbol{V})$ is the activation function of each unit:

$$o_i(\boldsymbol{V}) = \prod_{j=1}^{J} \exp\text{-}\frac{\left(V_j - \mu_{i,j}\right)^2}{2\sigma_{i,j}^2} \tag{3}$$

where V_j is the j-th element of input control vector $\boldsymbol{V}$. To train the network we have used a novel algorithm based on the Expectation-Maximization (EM) algorithm [5]. We can write the function to be fitted as $y(\boldsymbol{V}) = g(\boldsymbol{V}) + e$, where e is the error of the approximation, which can be assumed to be zero mean white Gaussian noise. Following the ideas expressed in [5,6], the observations $y(\boldsymbol{V})$ can be decomposed into its signal and noise components:

$$y_i(\boldsymbol{V}) = g_i(\boldsymbol{V}) + e_i \tag{4}$$

where the residuals e_i are obtained by decomposing the total error e into I components $e_i = t_i\, e$, and the decoupling variables t_i are restricted to sum the unity. Using this decomposition, the EM algorithm can be described as

E step : for $i = 1, \ldots, I$ compute $y_i(\boldsymbol{V}) = g_i(\boldsymbol{V}) + t_i\, e$

M step : for $i = 1, \ldots, I$ compute $\min\limits_{(\lambda_i;\mu_{i,j};\sigma_{i,j})} \sum_k \left(y_i(\boldsymbol{V}_k) - g_i(\boldsymbol{V}_k)\right)^2$

where k indexes the data points available. The M step is performed by means of a gradient-based method. By using this EM procedure the original complicated multiparameter optimization problem is decomposed into a set of more simple problems, which consists of estimating the parameters of each GRBF unit separately. This network provides a useful small-signal transistor model, which is able to reproduce the intermodulation distortion behavior. However, it has a clear local nature and when the dynamic voltages (v_{ds},v_{gs}) are large the Taylor series expansion has a loss of accuracy. In this case it is necessary to look for a large-signal model.

THE LARGE SIGNAL BEHAVIOUR

In the case of large dynamic signal, our modeling problem consists of obtaining a function G: $\Re^4 \rightarrow \Re$, which approximates the nonlinear mapping from the input space $\boldsymbol{V} = (V_{dso}, V_{gso}, v_{ds}, v_{gs})$ of static and dynamic pulsed voltages to the output space. A smoothed piecewise linear (SPWL) model [3] has been used, and is given by

$$I_{ds}^{LS)}(\boldsymbol{V}) = \boldsymbol{a} + \boldsymbol{BV} + \sum_{i=1}^{\theta} \boldsymbol{c}_i \frac{1}{\gamma} \ln(\cosh(\gamma(\langle \boldsymbol{\alpha}_i, \boldsymbol{V}\rangle - \beta_i))) \tag{5}$$

where $\boldsymbol{V}$ and $\boldsymbol{\alpha_i}$ are vectors of the same dimension, M, as the input space; $\boldsymbol{a}$ and $\boldsymbol{c_i}$ are vectors of the same dimension of the output space, N; $\boldsymbol{B}$ is an $N \times M$ matrix, β_i is a scalar, $<\,.\,>$ denotes the inner product and γ is another scalar that controls the smoothness of the model. This model is an extension of the well-known canonical piecewise linear model proposed by Chua [7], which smoothes the transition between linear regions by means of the function $f(x)=ln(cosh(\gamma x))/\gamma$. The training process consists of an iterative method that first moves the partition of the input space (given by $\boldsymbol{\alpha_i}$ and β_i) applying a gradient-based algorithm, and then estimates the optimal coefficients $\boldsymbol{a}$, $\boldsymbol{B}$ and $\boldsymbol{c_i}$. This model yields a smooth and derivable approximation with a low number of parameters and a reduced computational burden [3].

With this model we obtain an accurate model of the large-signal behavior of the device, but it fails when it is applied to small-signal analysis, because this model does not fit accurately enough up to the third order derivatives of the characteristic function of the device.

THE COMBINED NEURAL NETWORK

In order to provide a coherent single global model capable of representing the whole transistor behavior, a simple alternative could be to combine the large and small-signal modules into a single model as it is shown in Figure 1. The two networks are combined by means of a simple fuzzy combiner that weights each module taking into account the distance, d, of the instantaneous voltages with respect to the bias point

$$d = \sqrt{v_{ds}^2 + v_{gs}^2} \tag{6}$$

using this distance, the membership function for the small-signal regime is given by

$$\mu_{SS}(d) = \begin{cases} 1, & |d| \le d_1 \\ \dfrac{d_2 - d}{d_2 - d_1}, & d_1 < d < d_2 \\ 0, & |d| \ge d_2 \end{cases} \qquad (7)$$

whereas for the large-signal regime we have $\mu_{LS}(d) = 1 - \mu_{SS}(d)$. In (7), d_1 and d_2 are fixed parameters.

MODEL VALIDATION

We have applied the network described above to the modeling of a MESFET transistor. The data used to train and test the model were obtained from an analytical model [8] developed from a deep study of the particular behavior of a NE72084 MESFET. The large-signal module has been trained using 12 basis functions (hyperplanes), which implies 66 parameters. The smoothing parameter, γ, has been trained from the information of the derivative with respect to V_{gs}, because most of the nonlinear behavior occurs along that direction. The small-signal module has been trained using 8 basis functions (gausssians), which implies 114 parameters. The fuzzy combiner was set to make the linear transition between modules between 0.25 and 0.3 V of distance with respect to the bias point: that is, $d_1 = 0.25$ and $d_2 = 0.3$. Figure 2 shows the *I/V* characteristic function for a bias point of (V_{ds}= 3.5 V, V_{gs}=-1 V) and the approximation provided by the global model. In Figure 3 we show the behavior in a small-signal situation. Figure 3 a) shows the parameter G_m (derivative with respect to $V_{gs} = V_{gso} + v_{gs}$ at the bias point) as a function of the bias point, while Figure 3 b) shows the corresponding estimate given by the modular model. Figure 3 c) represents the approximation of the derivative provided by the large-signal module alone: the result obtained is clearly worse than that given by the modular model. Besides, it must be noticed that G_m is the first derivative; for higher order derivatives we observe a stronger degradation. Results obtained using other typical neural network architectures (with an equivalent number of parameters), such as a multilayer perceptron, suggest that a single network can not capture the information needed to accurately model both the large and small-signal behaviors.

CONCLUSIONS

A new neural network structure has been presented that performs a global modeling of microwave transistors. It provides an accurate approximation of the whole behavior of the device combining two modules that capture a different kind of behavior. A first module is responsible of capturing the nonlinear dynamic *I/V* characteristic of the device, which drives the large signal behavior. The other module is responsible of the local reconstruction of the *I/V* characteristic, taking into account the information of the derivatives, in order to represent the small-signal intermodulation behavior. The global model presents a reduced number of parameters, and the computational burden to carry out the training process is lower than that required by other networks, like the MLP, which allows an easy implementation in practical circuit simulators.

REFERENCES

[1] T. Fernández et al. "Extracting a bias-dependent large signal MESFET model from pulsed I/V measurements". *IEEE Trans. Microwave Theory Tech.*, vol 44, no. 3, pp. 372-378, 1996.

[2] A. M. Crosmun, S. Maas, "Minimization of intermodulation distortion in GaAs MESFET small-signal amplifiers", *IEEE Trans. Microwave Theory Tech.*, vol. 37, no. 9, pp. 1411-1417, 1989.

[3] M. Lázaro, I. Santamaría, C. Pantaleón et al. "Smoothing the canonical piecewise linear model: an efficient and derivable large-signal model for MESFET/HEMT transistors", Submitted to *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*.

[4] I. Santamaría, M. Lázaro, C. Pantaleón, J. García, A. Tazón, A. Mediavilla, "A nonlinear MESFET model for intermodulation analysis using a generalized radial basis function network", *Neurocomputing*, vol. 25, pp. 1-18, 1999.

[5] I. Santamaría, M. Lázaro, C. Pantaleón, "EM-based training of GRBF networks with application to nonlinear MESFET modeling" submitted to ICASSP2000.

[6] M. Feder, E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 477-489, 1988.

[7] L. O. Chua, A. C. Deng, "Canonical piecewise-linear modeling", *IEEE Trans. Circuits Syst.*, vol. 33, no. 5, pp. 511-525, 1986.

[8] C. Navarro et al, "Large signal dynamic properties of GaAs MESFET/HEMT devices under optical illumination", *Proc. of the GAAS'98 Symposium*, pp. 350-353, Amsterdam, 1998.

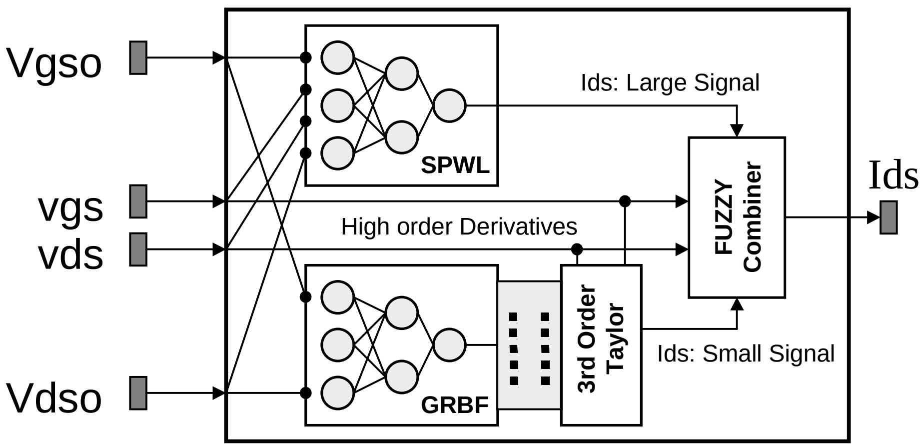

Fig.1 Modular Neural Network Structure

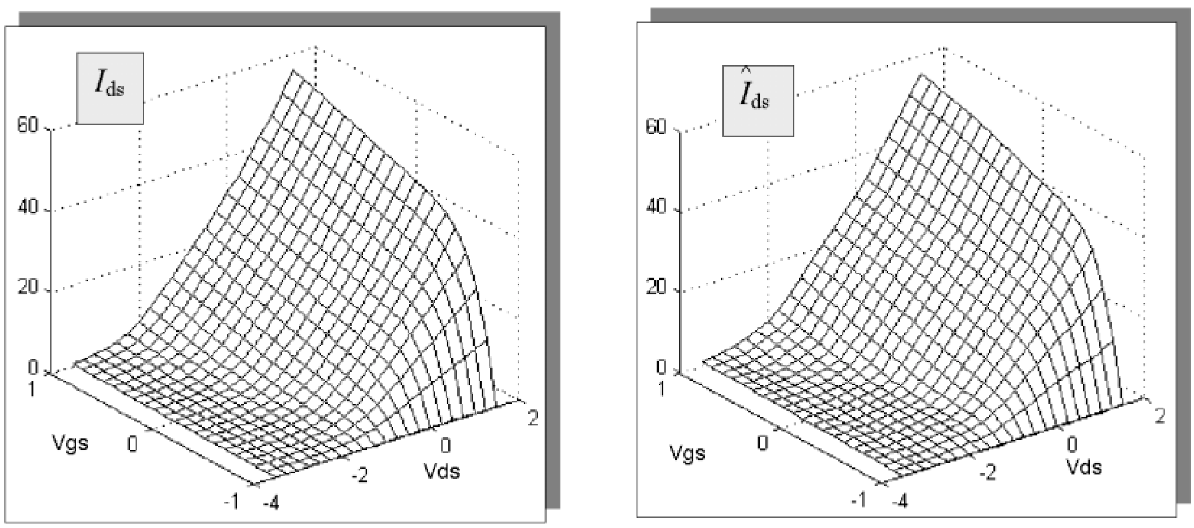

Fig.2: Characteristic *I/V* for a bias point (V_{ds}= 3.5 V, V_{gs}=-1 V). Original (left) and approximation (right).

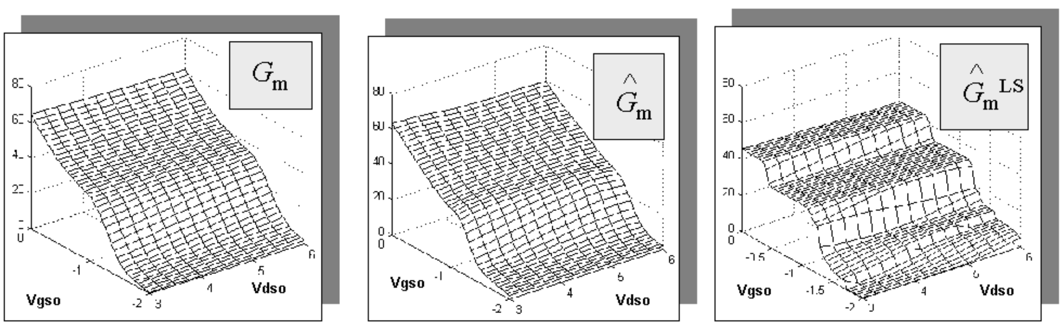

Fig.3: Parameter Gm, a) original, b) proposed modular model, c) Only LS model