ELSEVIER

Neurocomputing 25 (1999) 1–18

NEUROCOMPUTING

A nonlinear MESFET model for intermodulation analysis using a generalized radial basis function network[1]

Ignacio Santamaría*, Marcelino Lázaro, Carlos J. Pantaleón, Jose A. García, Antonio Tazón, Angel Mediavilla

DICOM, ETSII y Telecomunicación, Univ. Cantabria, Av. Los Castros s/n, 39005 Santander, Spain

Received 27 May 1998; accepted 24 November 1998

Abstract

In this paper we use a generalized radial basis function (GRBF) network to model the intermodulation properties of microwave GaAs MESFET transistors under dynamic operation. The proposed model receives as input the bias voltages of the transistor and provides as output the derivatives of the drain-to-source current, which are responsible for the intermodulation properties. The GRBF network is a generalization of the RBF network, which allows different variances for each dimension of the input space. This modification allows to take advantage of the soft nonlinear dependence of the output derivatives with the drain-to-source bias voltage. The learning algorithm chooses the GRBF centers one by one in order to minimize the output error. After selecting each new center from the training set, the centers and variances of the global network are optimized by applying gradient descent techniques. Finally, the amplitudes are obtained by solving a least-squares problem. The effectiveness of the proposed GRBF model is validated through load-pull intermodulation prediction based on the experimental nonlinear characterization of an NE72084 MESFET device.

Keywords: Generalized radial basis function (GRBF); Radial basis function (RBF); Nonlinear modeling; MESFET modeling; Gradient descent learning

1. Introduction

The design of microwave and millimeter-wave circuits requires accurate modeling of the nonlinear behavior of active devices such as a metal semiconductor field effect

* Corresponding author. Tel.: + 34 42 201552; fax: + 34 42 201488; e-mail: nacho@gtas.dicom.unican.es

[1] This work has been supported by CYCIT grant TIC96-0500-C10-07.

0925-2312/99/$ – see front matter
PII: S0925-2312(98)00106-4

transistor (MESFET). Specifically, in multiple carrier systems the MESFETs nonlinear behavior causes intermodulation distortion. The prediction and modeling of the intermodulation effects is an important issue for the performance of broadband communication systems.

As it is shown in [3], to predict the MESFETs intermodulation behavior, it is necessary to approximate not only its current/voltage (I/V) nonlinear characteristic, but also its derivatives. However, the conventional nonlinear techniques applied to model the MESFETs characteristics, such as the widely employed analytical functions [4,9], the canonical piecewise-linear model [2], or the use of look-up tables [14], fail to fit simultaneously the input–output nonlinear function and its higher-order derivatives.

Recently, some attempts have been made to model the nonlinear behavior of active devices and circuits by using neural networks [15,16,18]. Neural networks have the capability of approximating any nonlinear function and the ability to learn from experimental data. These characteristics make neural networks a good alternative to overcome some of the drawbacks of the traditional modeling techniques. However, practically all of these neural approaches only consider the use of the multilayer perceptron (MLP), and most of them do not have the capability to cope with intermodulation effects.

In this paper, we propose to use a generalized radial basis function (GRBF) network to model the derivatives of the I/V nonlinear characteristic of a MESFET. This network receives as input the bias voltages and produces at its output the coefficients of a truncated two-dimensional Taylor series. These parameters can be used to model the nonlinear drain current of the MESFET in small-signal regime.

The GRBF network is a generalization of the RBF network, which allows different variances for each dimension of the input space. By replacing the radial Gaussian kernels with elliptical basis functions we take advantage of the soft nonlinear dependence of the network's output with the drain-to-source bias voltage. A GRBF neuron responds to a larger region of the input space than a conventional fixed-variance RBF neuron. Therefore, when compared to the RBF network, the GRBF network reduces drastically the number of units required to obtain an accurate model.

The rest of the paper is organized as follows. Section 2 reviews some aspects of nonlinear MESFET modeling with special emphasis on the intermodulation behavior. Section 3 describes the GRBF network and discusses its advantages for this particular problem. In Section 4, we describe the learning algorithm for the network parameters: centers, variances and amplitudes. Simulation results using experimental data are presented in Section 5 to demonstrate the effectiveness of the proposed approach. Finally, in Section 6 the main conclusions are summarized.

2. Nonlinear behavior of the MESFET

Fig. 1 represents the most widely accepted equivalent nonlinear circuit of a MESFET in its saturated region. The predominant nonlinear element is the drain-to-source current I_{ds}, which depends on both the drain-to-source, V_{ds}, and gate-to-source, V_{gs},

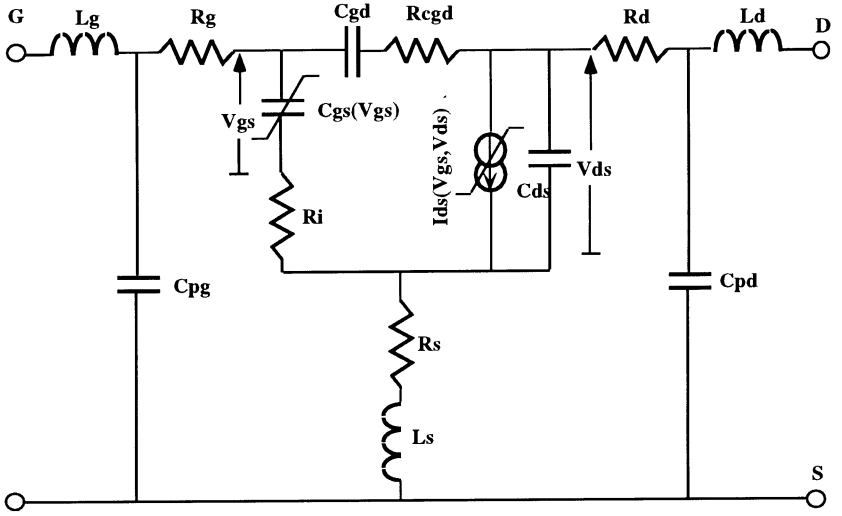


Fig. 1. Nonlinear equivalent circuit of a MESFET.

bias voltages: this dependence is denoted as the I/V characteristic. As it is shown in [8], the nth-order intermodulation output power varies fundamentally as the square of the nth derivative of the I/V characteristic. Usually, for mixers and amplifiers it is necessary to model up to the third intermodulation product; therefore, our model must accurately fit up to the third derivative of the mentioned nonlinearity in the operating region of our concern.

When we apply a small-signal RF input around a bias point, the drain current I_{ds} depends on the bias point (V_{ds},V_{gs}) and the instantaneous small-signal voltages (v_{ds},v_{gs}). In this case, I_{ds} can be represented in a small interval around the bias point by the following two-dimensional truncated Taylor series expansion:

$$I_{ds} = I_{ds0} + G_m v_{gs} + G_d v_{ds} + G_{m2} v_{gs}^2 + G_{md} v_{gs} v_{ds} + G_{d2} v_{ds}^2 G_{m3} v_{gs}^3 + G_{m2d} v_{gs}^2 v_{ds} + G_{md2} v_{gs} v_{ds}^2 + G_{d3} v_{ds}^3 , \quad (1)$$

where I_{ds0} is the dc drain current, v_{ds} and v_{gs} are the incremental drain-to-source and gate-to-source voltages, respectively; and ($G_m, \ldots, G_{d3}$) are coefficients related to the nth-order derivatives of the I/V characteristic evaluated at the bias point. For instance, G_{md2} is given by

$$G_{md2} = \frac{1}{2} \frac{\partial^3 I_{ds}}{\partial V_{gs} \partial V_{ds}^2} . \quad (2)$$

The parameters ($G_m, \ldots, G_{d3}$) can be determined for a number of bias voltages from intermodulation power measurements for two tones excitation [12]: they are the output targets of our neural network model. On the other hand, the input patterns are the drain-to-source and gate-to-source bias voltages.

After obtaining a set of real measurements using the procedure described in [12], our modeling problem can be stated as the approximation of a multidimensional function. Given a set of input–output patterns, we have to obtain a function (model) $\boldsymbol{G}: \mathscr{R}^2 \rightarrow \mathscr{R}^{10}$ that approximates the nonlinear mapping from the input space of bias voltages $\boldsymbol{V} = (V_{ds}, V_{gs})$, to the output space of model parameters $\boldsymbol{G}(\boldsymbol{V}) = (I_{ds0}, G_m, G_d, G_{m2}, G_{md}, G_{d2}, G_{m3}, G_{m2d}, G_{md2}, G_{d3})$.

Once the neural network is trained, it provides, for each input bias point, a set of 10 parameters which can be used to reconstruct a small-signal bias-dependent MESFET model by using the truncated Taylor series expansion given by Eq. (1). In this way we assure an accurate prediction of intermodulation effects around the bias point.

3. Generalized radial basis function network

The nonlinear input–output mapping between the bias voltages and the model parameters can be approximated using conventional feedforward layered networks such as the radial basis function (RBF) network or the multilayer perceptron (MLP). The MLP and the RBF network are both universal approximators, as shown in [5,11]. However, these two networks differ from each other in two important respects: first, an RBF unit using a Gaussian kernel performs a local approximation, while an MLP constructs a global approximation to the input–output mapping; this leads to larger size RBF networks (this fact becomes a serious drawback in high dimensional input spaces). On the other hand, the linear characteristic of the output layer of the RBF networks leads to a faster training.

In this paper, we concentrate on the application of RBF networks. Specifically, we consider an extension of the RBF network which allows a different variance for each input dimension. The relaxation of the radial constraint transforms the standard Gaussian kernels with circular symmetry into elliptic basis kernels, which can reduce the dimensionality of the input space. This scheme is denoted as generalized radial basis function (GRBF) network [7].

The idea of constructing hyperellipses around the centers of the basis functions was considered in [13] in the context of regularization theory. A generalization of the elliptic kernel denoted as "Gaussian bar" unit was proposed in [6] to improve the performance of the RBF network in the presence of irrelevant inputs. This generalization sums the weighted Gaussian responses along each input dimension, while the conventional RBF and the GRBF obtains a nonweighted product.

In order to understand the rationale for using an GRBF network, Fig. 2a and b show the measured coefficients G_{md} and G_{m3} as a function of the bias point $\boldsymbol{V} = (V_{ds}, V_{gs})$, respectively. The shape of G_{md} and G_{m3} along the V_{gs} axis suggests that they could be approximated by a combination of Gaussians (a single Gaussian in the case of G_{md}). However, they have a quasi-linear dependence with V_{ds}. In order to adjust a neural model with a minimum number of parameters, the activation functions in the hidden layer should respond to a localized region along V_{gs} and to a non-localized region along V_{ds}. By allowing a different variance for each input dimension, the GRBF network attains this kind of semilocal behavior. On the other hand, by

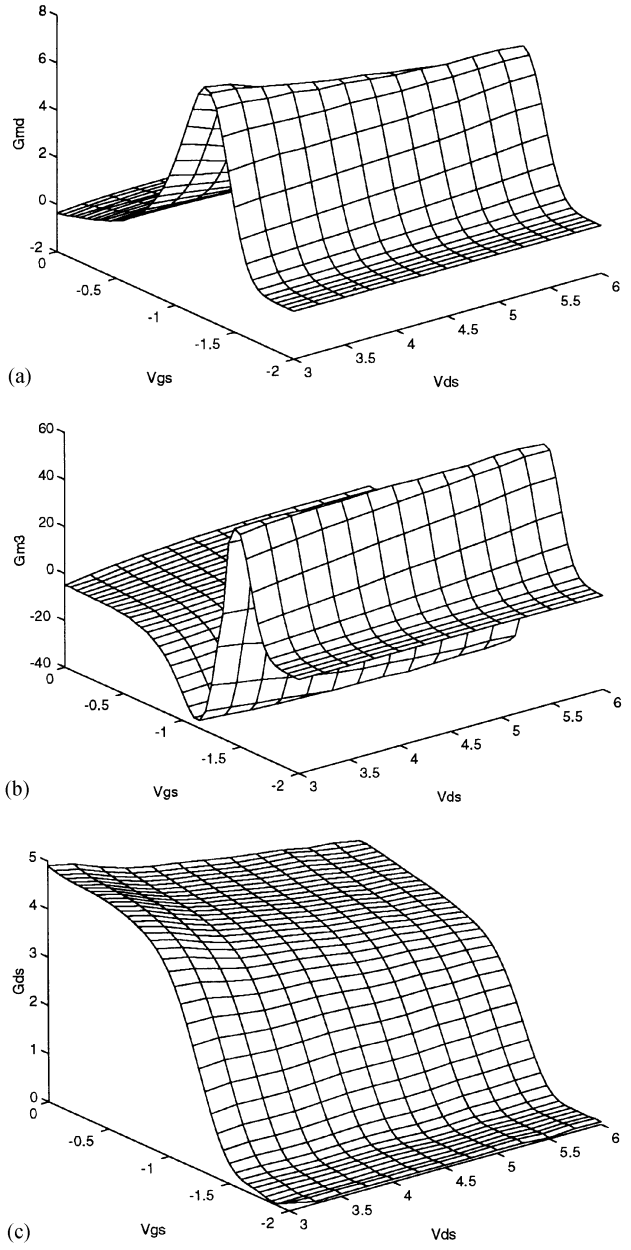


Fig. 2. (a) Measured G_{md} coefficient, (b) G_{m3} and (c) G_{ds}, as a function of the bias voltage (V_{ds},V_{gs}).

broadening the variance along both directions, an GRBF unit is able to respond to a large region of the input space: this is appropriate to approximate some parameters of the model that do not have a Gaussian shape along the V_{gs} axis, such as G_{ds}, (see Fig. 2c). Therefore, the GRBF lies somewhere between the highly local RBF and the global MLP.

For notation simplicity, let us decompose the global mapping performed by the GRBF network ($\boldsymbol{G}: \mathscr{R}^{J} \rightarrow \mathscr{R}^{M}$) into a set of single-output networks as follows:

$$\boldsymbol{G}(\boldsymbol{V}) = (g_1(\boldsymbol{V}), \ldots, g_M(\boldsymbol{V})), \tag{3}$$

each scalar output is given by

$$g_k(\boldsymbol{V}) = \sum_i \lambda_{ik} \prod_j \exp - \frac{(V_j - \mu_{ij})^2}{2\sigma_{ij}^2}, \quad k = 1, \ldots, M, \tag{4}$$

where i indexes the GRBF units, j the input dimensions and k the output dimensions.

The GRBF can be viewed as an RBF for which the Euclidean norm is replaced by a weighted norm. Specifically, Eq. (4) can be rewritten as

$$g_k(\boldsymbol{V}) = \sum_i \lambda_{ik} \prod_j \exp - \|\boldsymbol{V} - \mu_i\|_{\boldsymbol{W}_i}^2, \tag{5}$$

where

$$\|\boldsymbol{V} - \mu_i\|_{\boldsymbol{W}_i}^2 = (\boldsymbol{V} - \mu_i)^{\mathrm{T}} \boldsymbol{W}_i^{\mathrm{T}} \boldsymbol{W}_i (\boldsymbol{V} - \mu_i) \tag{6}$$

$\mu_i = (\mu_{i1}, \mu_{i2})^{\mathrm{T}}$ and $\boldsymbol{W}_i$ is a diagonal matrix given by

$$\boldsymbol{W}_i = \frac{1}{\sqrt{2}} \mathrm{diag}\left(\frac{1}{\sigma_{i1}}, \frac{1}{\sigma_{i2}}\right). \tag{7}$$

In its more general form, the GRBF network considers a nondiagonal norm weighting matrix $\boldsymbol{W}_i$ [7]; this results in elliptic basis kernels which have adjustable orientations. In this paper, however, we consider a simplified version of the GRBF network which uses a diagonal weighting matrix $\boldsymbol{W}_i$: i.e., we allow the variances to vary along each input dimension, but we do not allow the elliptic kernels to rotate.

For comparison purposes, the "Gaussian bar" activation function proposed in [6] is given by

$$g_k(\boldsymbol{V}) = \sum_i \sum_j \lambda_{ijk} \exp - \frac{(V_j - \mu_{ij})^2}{2\sigma_{ij}^2}, \quad k = 1, \ldots, M. \tag{8}$$

By adding the weighted Gaussian responses along each input direction as in Eq. (8), we can get an additional improvement at the expense of an increase in the number of parameters: the number of amplitudes of the output layer is multiplied by the dimension of the input space. In our simulations, however, for a given number of model parameters the GRBF network outperforms the "Gaussian bar" network. For the particular application considered in this paper, where the objective is to obtain an accurate model with as few parameters as possible, we have observed that most of the

improvement comes from using a different variance along each direction, while the use of a different amplitude along each direction does not result in an additional (noticeable) improvement.

4. Network optimization

For RBF networks, the simplest learning strategy consists of fixing the centers and the variances defining the nonlinearities in the hidden layer; and then obtaining the amplitudes of the output layer by solving a linear least-squares problem. The centers can be selected at random, by applying vector quantization algorithms to the input data set [10], or by applying more efficient strategies such as the orthogonal least-squares (OLS) method [1]. On the other hand, the variance is fixed for all the units and is selected according to the spread of the centers.

For the particular case of GRBF networks some modifications of the learning strategy are required in order to obtain an efficient and parsimonious model. First, the variances along each direction and the centers are obtained by applying a gradient descent algorithm. The error function is given by

$$E = \sum_p \sum_k (y_k(\boldsymbol{V}_p) - g_k(\boldsymbol{V}_p))^2 \tag{9}$$

and the gradient equations for the variances and centers are

$$\frac{\partial E}{\partial \sigma_{ij}} = -2 \sum_p \sum_k e_k(\boldsymbol{V}_p) o_i(\boldsymbol{V}_p) \lambda_{ik} \frac{1}{\sigma_{ij}} \left(\frac{V_{pj} - \mu_{ij}}{\sigma_{ij}} \right)^2, \tag{10}$$

$$\frac{\partial E}{\partial \mu_{ij}} = -2 \sum_p \sum_k e_k(\boldsymbol{V}_p) o_i(\boldsymbol{V}_p) \lambda_{ik} \frac{1}{\sigma_{ij}} \left(\frac{V_{pj} - \mu_{ij}}{\sigma_{ij}} \right), \tag{11}$$

where p indexes the input patterns, k the output dimensions, $\boldsymbol{V}_p$ is the pth input pattern, $y_k(\boldsymbol{V}_p)$ is the desired (measured) output, $g_k(\boldsymbol{V}_p)$ is the output of the network (4), $e_k(\boldsymbol{V}_p) = y_k(\boldsymbol{V}_p) - g_k(\boldsymbol{V}_p)$ is the network error and $o_i(\boldsymbol{V}_p)$ is the output of neuron i

$$o_i(\boldsymbol{V}_p) = \prod_j \exp - \frac{(V_{pj} - \mu_{ij})^2}{2\sigma_{ij}^2}. \tag{12}$$

The second modification consists of selecting the GRBF units one by one until the squared error decreases below some threshold or a maximum number of units is reached. Looking again at Fig. 2a we can understand the benefits of the proposed procedure: the G_{md} parameter can be approximated by placing a single GRBF unit and broadening its variance along the drain-to-source input voltage direction. Therefore, the selection of the GRBF units one by one allows to take full advantage of their semilocal approximation capabilities.

At each step, the new GRBF center is selected from the input training set in order to get a maximum decrease in the squared output error. If the number of input data patterns is small, an exhaustive search can be applied; otherwise an "ad hoc" technique should be applied to reduce the computational burden associated to the search. For instance, a suitable procedure is to search only in a neighborhood around the point of maximum error.

After determining the initial parameters of the new GRBF unit, the centers and variances of the global GRBF network are updated according to Eqs. (10) and (11) and, finally, the amplitudes of the global network are reestimated by solving the following linear least-squares problem:

$$\begin{pmatrix} o_1(\boldsymbol{V_1}) & \cdots & o_N(\boldsymbol{V_1}) \\ \vdots & \ddots & \vdots \\ o_1(\boldsymbol{V_P}) & \cdots & o_N(\boldsymbol{V_P}) \end{pmatrix} \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{M1} \\ \vdots & \ddots & \vdots \\ \lambda_{1N} & \cdots & \lambda_{MN} \end{pmatrix} = \begin{pmatrix} y_1(\boldsymbol{V_1}) & \cdots & y_M(\boldsymbol{V_1}) \\ \vdots & \ddots & \vdots \\ y_1(\boldsymbol{V_P}) & \cdots & y_M(\boldsymbol{V_P}) \end{pmatrix}, \tag{13}$$

where $o_i(\boldsymbol{V}_p)$ was defined in Eq. (12), $\boldsymbol{Y}(\boldsymbol{V}_p) = (y_1(\boldsymbol{V}_p), \ldots, y_M(\boldsymbol{V}_p))$ is the pth target output, and λ_{ik} is the amplitude connecting the ith GRBF unit in the hidden layer to the kth unit in the output layer.

Finally, the proposed algorithm can be summarized in the following steps:

1. Initialize the learning parameters for the variances and the centers η_v, η_c, respectively; the final number of neurons N, and the variance of the GRBF units according to the input data spread.
2. Initialize the output of the model $\boldsymbol{G}^0(\boldsymbol{V}_p) = (0, \ldots, 0)$, for $p = 1, \ldots, P$; where $\boldsymbol{G}^t(\boldsymbol{V}_p)$ denotes the output of the network after adding the tth neuron.
3. For $t = 1$–N
   3.1. Obtain the output error for the model with $t-1$ neurons

$$\boldsymbol{E}^t(\boldsymbol{V}_p) = \boldsymbol{Y}(\boldsymbol{V}_p) - \boldsymbol{G}^{t-1}(\boldsymbol{V}_p), \quad p = 1, \ldots, P.$$

   3.2. Obtain the new GRBF unit which most reduces the error function $\boldsymbol{E}^t$.
   3.3. Update the variances and centers of all the neurons by applying a gradient descent algorithm

$$\sigma_{ij}^{(s+1)} = \sigma_{ij}^{(s)} - \eta_v \frac{\partial E}{\partial \sigma_{ij}} \quad \text{for } i = 1, \ldots, t,$$

$$\mu_{ij}^{(s+1)} = \mu_{ij}^{(s)} - \eta_c \frac{\partial E}{\partial \mu_{ij}} \quad \text{for } i = 1, \ldots, t,$$

   where s denotes iteration.

3.4. Estimate the amplitudes of the global GRBF network according to Eq. (13).
3.5. Obtain the new output of the GRBF network with t neurons:

$$\boldsymbol{G}^t(\boldsymbol{V}_p) \quad \text{for } p = 1, \ldots, P$$

end.

Let us remark that, in Step 3.3, we minimize the global error function given by Eq. (9), therefore we use the gradients (10) and (11). Moreover, we update the centers and variances of all the neurons and not only the new neuron located at time t. Finally, the iterations at this step are carried out until the error decreases below some threshold or until a maximum number of iterations is reached.

5. Simulation results

The proposed GRBF network was used to model a microwave NE72084 MESFET. The model outputs (I_{ds0},G_m,G_d,G_{m2},G_{md},G_{d2},G_{m3},G_{m2d}, G_{md2},G_{d3}) were measured at different bias voltages. Specifically, V_{ds} was swept from 3 to 6 V in steps of 0.25 V, and V_{gs} from -2 to 0 V in steps of 0.05 V; thus giving a total of 533 input–output training patterns.

In our experience, the partitioning of this particular data set into a training set and a testing set does not improve necessarily the network's generalization performance. This could be due to the fact that we have a small number of input measurements and, moreover, these measurements have a low level of noise (for most of the parameters). Therefore, we choose to use the whole measurement data set for training.

In this section we compare the performance of three different neural network models: the RBF network, the MLP and the proposed GRBF network. The objective is to obtain an accurate model with a small number of parameters (parsimonious). For this particular application, in order to implement the nonlinear mapping $\boldsymbol{G}: \mathscr{R}^2 \to \mathscr{R}^{10}$, an GRBF network with N neurons in the hidden layer (denoted as GRBF(N)) requires $14N$ parameters. For each new GRBF neuron we have to specify its center (2 parameters), its variance (2 parameters) and 10 amplitudes connecting the hidden layer with the output. On the other hand, an MLP(N) with one hidden layer and an RBF(N) require $14N$ and $13N$ parameters, respectively.

The accuracy of each model was measured in terms of the SNR for each scalar output. For instance, for G_{m3} the SNR is evaluated as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum_p G_{m3,p}^2}{\sum_p (G_{m3,p} - \hat{G}_{m3,p})^2}\right), \tag{14}$$

where p indexes the training patterns, $G_{m3,p}^2$ is the desired (real measurement) output and $\hat{G}_{m3,p}^2$ is the output estimated by the network.

A significant effort was made to optimize the algorithm parameter settings (learning rates, variances for the RBF network, etc). In particular, we used the following

parameters:

- MLP(N): MLP with N neurons in the hidden layer trained with a backpropagation algorithm using an adaptive learning rate (initial learning rate $\eta = 0.0001$). Number of training epochs = 5000.
- RBF(N): RBF network with N units. The RBF centers are selected using the OLS algorithm [1]. We used a fixed variance $\sigma^2 = 0.125$.
- GRBF(N): GRBF network with N units. We used the algorithm described in Section 4 with the following parameters: $\eta_v = 2e-4$, $\eta_c = 1e-3$; maximum number of iterations for updating the variances and centers = 500; initial variance $\sigma^2 = (0.025, 0.1)$.

Figs. 3–5 show the output of an RBF(10), an MLP(8) and the proposed GRBF(8) model, for the output parameters G_{md}, G_{m3} and G_{ds}, respectively. These figures should be compared with the measured ones depicted in Fig. 2. The MLP and the GRBF model seem to capture the nonlinear behavior of these parameters. On the other hand, for the same number of model parameters, the highly local nature of the RBF network results in an excessively hilly output function. The values provided in Table 1 allow a more detailed comparison of the three models. We can point out the following conclusions: first, for a small number of model parameters the GRBF network provides the best results; secondly, to achieve a similar performance the MLP requires approximately twice as many parameters as the GRBF; finally, an RBF network requires the largest number of parameters.

Finally, in order to verify the validity of the proposed network to model the device nonlinear distortion behavior due to the drain current source, some output power and carrier to intermodulation (C/I) calculations were made using both the experimentally extracted derivatives and the GRBF network. As it is shown in [12,17], most of the existing models fail to reproduce the C/I behavior with varying load condition, and consequently the existence of optimum load values for low distortion designs. This shortcoming is due to their incapacity of fitting not only the predominant third derivative, G_{m3}, but also the cross third-order terms.

We made a typical two tones analysis with frequencies of 10 and 10.01 GHz and power levels slightly below the 1 dB compression point, assuring the small signal regime where the nonlinear distortion behavior is generally poorly predicted. We selected a commonly employed bias point for Class A amplifier applications, $V_{gs} = -0.2$ V and $V_{ds} = 3$ V. Fig. 6a and b show on a Smith chart the load-pull contours for the output power of the 10 GHz signal using the measured and the modeled (GRBF network) derivatives, respectively. The almost perfect agreement verifies the good network fitting of the first-order coefficients, G_{m1} and G_{ds}.

In Fig. 7a and b we show the contours for the C/I ratio indicating the power level differences between one of the referred signals (f_1 or f_2) and the undesired adjacent intermodulation products ($2f_1 - f_2$ or $2f_2 - f_1$). The prediction is significantly accurate, supporting a reliable reproduction of the second- and third-order coefficients not previously reported with the traditional techniques.

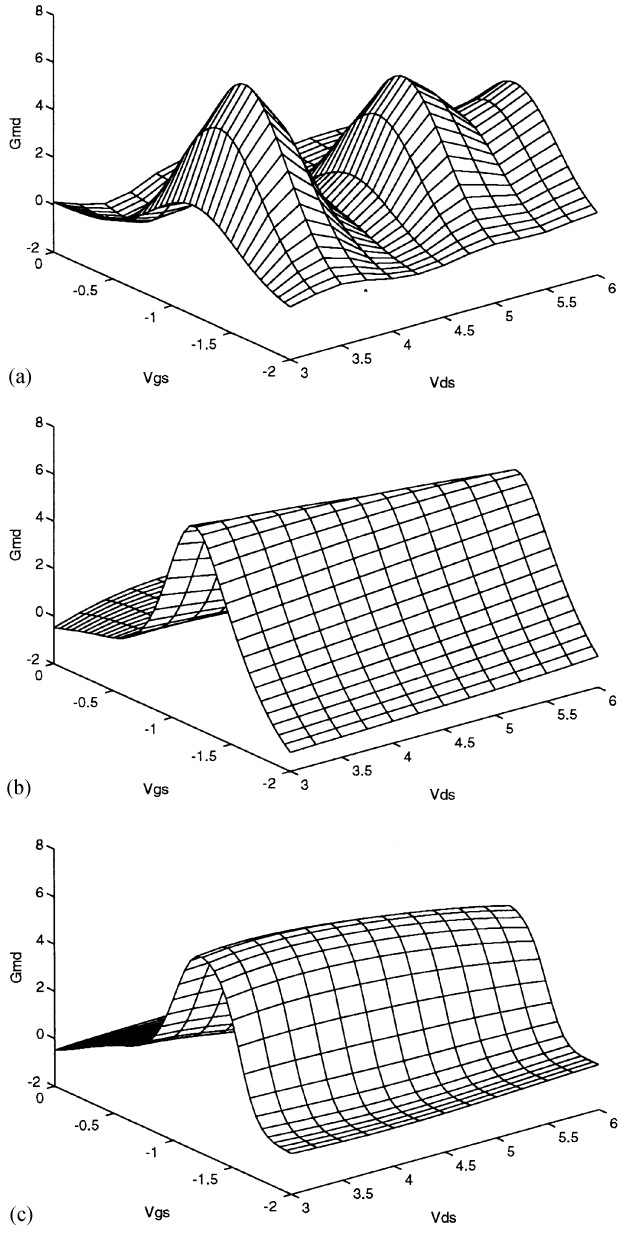


Fig. 3. (a) G_{md} estimate obtained by an RBF(10) network, (b) an MLP(8) and (c) an GRBF(8).

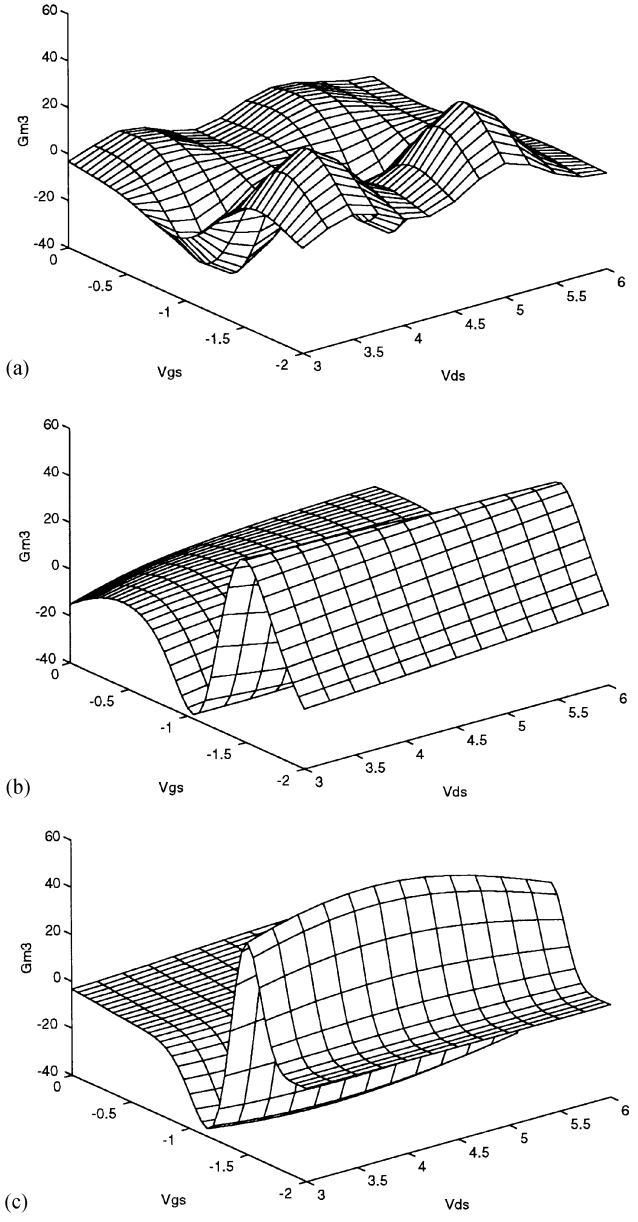


Fig. 4. (a) G_{m3} estimate obtained by an RBF(10) network, (b) an MLP(8) and (c) an GRBF(8).

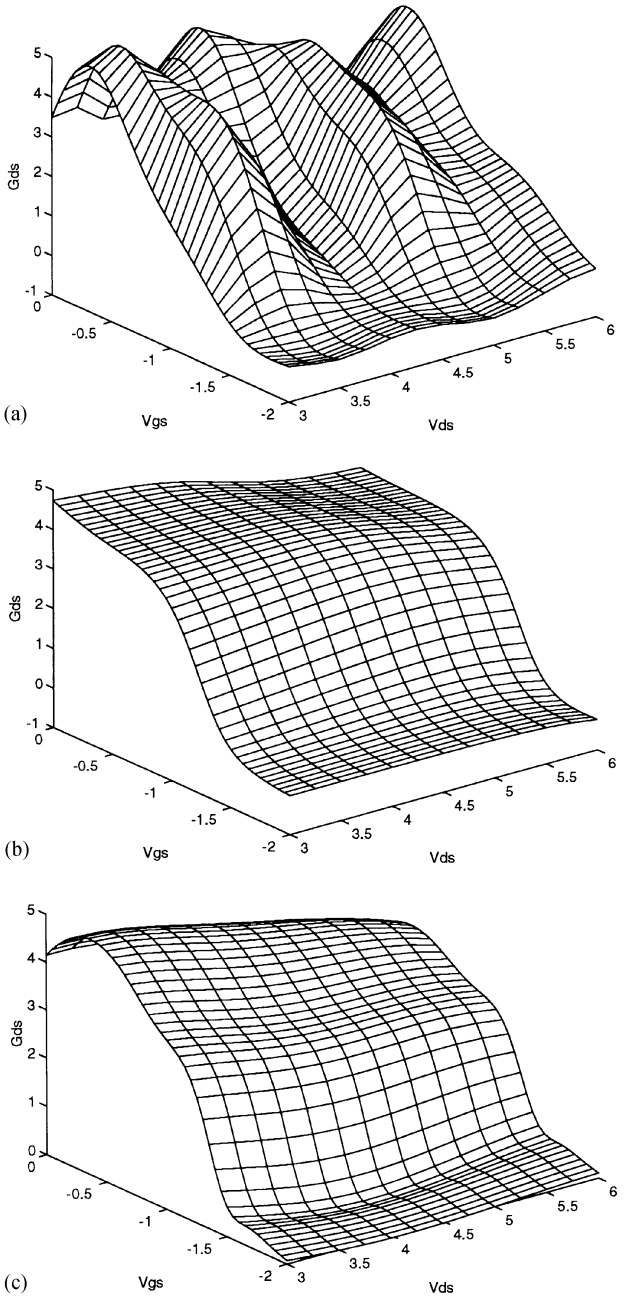


Fig. 5. (a) G_{ds} estimate obtained by an RBF(10) network, (b) an MLP(8) and (c) an GRBF(8).

Table 1
Comparison of the results obtained by the following neural models: GRBF(8), RBF(10), RBF(35), MLP(8), MLP(16). The first column indicates the number of parameters for each model, the following 10 columns show the SNR in dB for each output of the model

	N_{par}	I_{ds}	G_{d2}	G_{d3}	G_{ds}	G_{m1}	G_{m2}	G_{m2d}	G_{m3}	G_{md}	G_{md2}
GRBF(8)	112	22.8	19.1	15.7	25.4	27.2	17.0	18.0	18.5	17.9	16.1
RBF(10)	130	12.0	10.7	9.5	12.3	12.0	6.4	3.5	2.3	6.3	3.6
RBF(35)	455	21.5	18.6	16.7	23.3	23.0	17.3	11.4	10.4	17.3	12.2
MLP(8)	112	18.7	13.5	14.0	24.6	30.0	13.5	9.9	8.9	13.6	10.7
MLP(16)	224	20.0	14.4	16.7	29.1	31.7	16.3	15.4	14.3	16.3	14.2

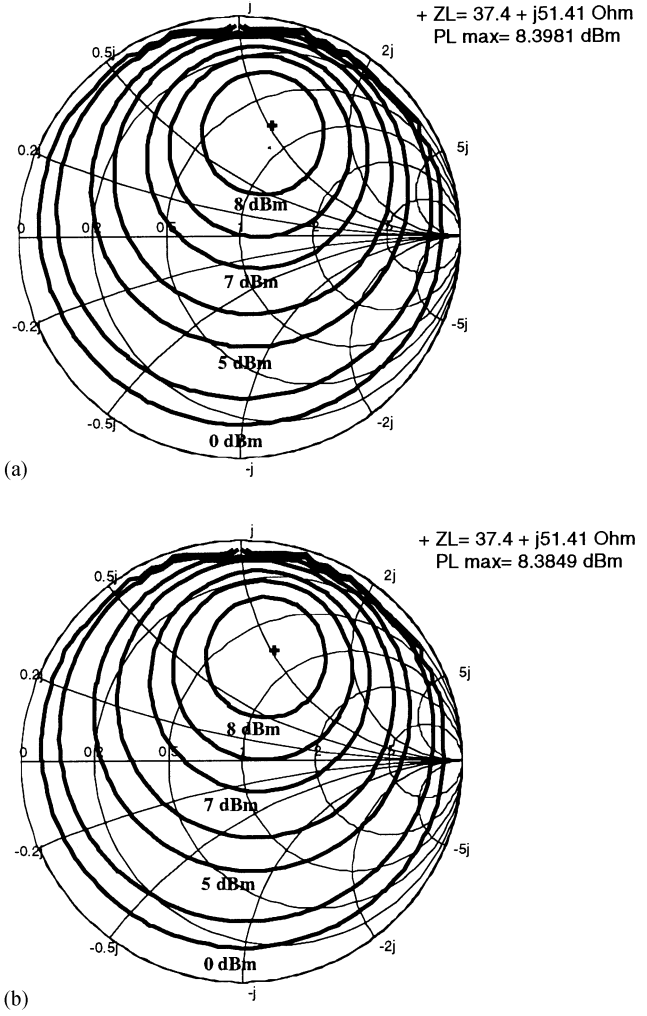


Fig. 6. Output power level of the NE72084 MESFET (a) using the measured parameters, (b) using the GRBF(8) network model.

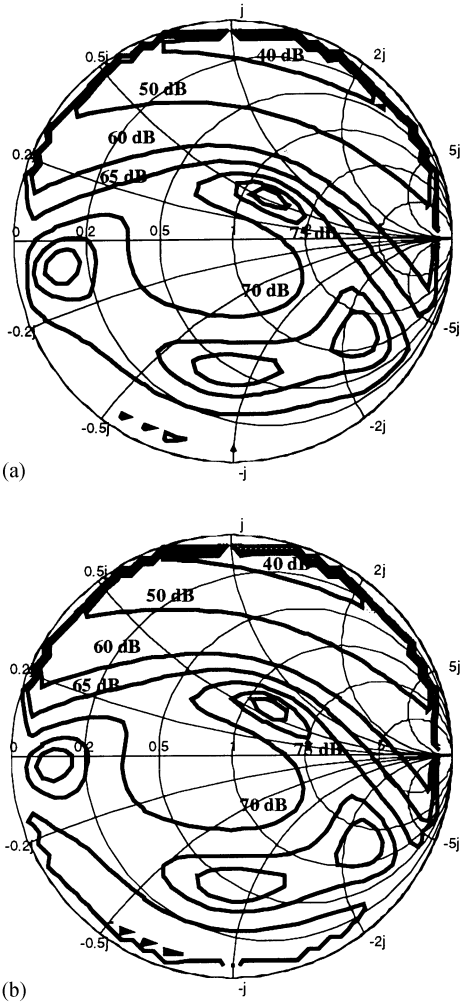


Fig. 7. Contours for the C/I ratio of the NE72084 MESFET (a) using the measured parameters, (b) using the GRBF(8) network model.

6. Conclusions

This paper has presented a new nonlinear MESFET model for intermodulation analysis using a generalized radial basis function (GRBF) network. By relaxing the radial constraint imposed by conventional RBF Gaussian kernels, the GRBF network yields a parsimonious and accurate model which can be used to simulate the small-signal bias-dependent MESFETs behavior.

An alternative learning procedure has been developed for the GRBF network: at each step, a new GRBF center is selected from the input training set in order to

get a maximum decrease in the squared output error. This suitable selection of the GRBF units one by one, combined with gradient descent techniques for updating the centers and variances, allows to take full advantage of their semilocal approximation capabilities.

When compared to the RBF network, the GRBF network reduces drastically the number of units required to obtain an accurate model. Moreover, using experimental measurements of an NE72084 MESFET, it has been shown that an MLP requires a higher number of parameters than the GRBF network to achieve a similar performance. The predicted load-pull behavior in output power and C/I ratio confirms the accuracy of the proposed GRBF network.

However, for some derivatives of the I/V characteristic, which have a nonlocalized support along the V_{gs} axis, the MLP provides a better extrapolation than the GRBF network. For this reason, the combination of a global approximator such as the MLP with the proposed GRBF network seems a promising line for further research.

Acknowledgements

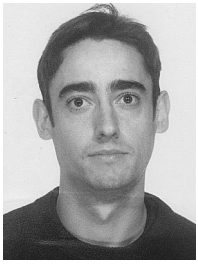
The authors would like to acknowledge the financial support of CYCIT grant TIC96-0500-C10-07. We would also like to thank the reviewers for their valuable suggestions and comments.

References

[1] S. Chen, C.F.N. Cowan, P.M. Grant, Orthogonal least-squares learning algorithm for radial basis functions networks, IEEE Trans. Neural Networks 2 (1991) 302–309.
[2] L.O. Chua, A.C. Deng, Canonical piecewise-linear modeling, IEEE Trans. Circuits Systems 33 (5) (1986) 511–525.
[3] A.M. Crosmun, S.A. Maas, Minimization of intermodulation distortion in GaAs MESFET small-signal amplifiers, IEEE Trans. Microwave Theory Technol. 37 (9) (1989) 1411–1417.
[4] W.R. Curtice, M. Ettenberg, A nonlinear GaAs FET model for use in the design of output circuits for power amplifiers, IEEE Trans. Microwave Theory Technol. 33 (12) (1985) 1383–1394.
[5] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Systems 2 (1989) 303–314.
[6] E. Hartman, J.D. Keeler, Predicting the future: advantages of semilocal units, Neural Comput. 3 (1991) 566–578.
[7] S. Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, New York, 1994.
[8] S.A. Maas, D. Neilson, Modeling MESFETs for intermodulation analysis of mixers and amplifiers, IEEE Trans. Microwave Theory Technol. 38 (12) (1990) 1964–1971.
[9] A. McCamant, G. McCormak, D. Smith, An improved GaAs MESFET for SPICE, IEEE Trans. Microwave Theory Technol. 38 (6) (1990) 822–824.
[10] J.E. Moody, C.J. Darken, Fast learning in networks of locally-tuned processing units, Neural Comput. 1 (1989) 281–294.
[11] J. Park, I.W. Sandberg, Universal approximation using radial-basis-function networks, Neural Comput. 3 (1991) 246–257.
[12] J.C. Pedro, J. Pérez, Accurate simulation of GaAs MESFETs intermodulation distortion using a new drain-source current model, IEEE Trans. Microwave Theory Technol. 42 (1) (1994) 25–33.

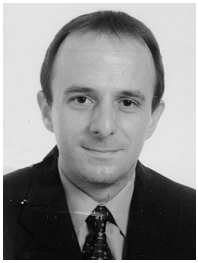
[13] T. Poggio, F. Girosi, Networks for approximation and learning, Proc. IEEE 78 (9) (1990) 1481–1497.
[14] D.E. Root, S. Fan, J. Meyer, Technology independent large signal nonquasi-static FET models by direct construction from automatically characterized device data, Proc. 21st European Microwave Conf., Stuttgart, Germany, 1991, pp. 923–927.
[15] J. Rousset, Y. Harkouss, J.M. Collantes, M. Campovecchio, An accurate neural network model of FET for intermodulation and power analysis, Proc. 26th European Microwave Conf., Prague, Czechoslovakia, 1996.
[16] K. Shirakawa, M. Shimiz, N. Okubo, Y. Daido, A large signal characterization of an HEMT using a multilayered neural network, IEEE Trans. Microwave Theory Technol. 45 (9) (1997) 1630–1633.
[17] D.R. Webster, D.G. Haigh, G. Passiopoulos, A.E. Parker, Distortion in short channel FET circuits, in: G. Machado (Ed.), Low-Power HF Microelectronics, IEE, London, 1996.
[18] A.H. Zaabab, Q.J. Zhang, M. Nakhla, A neural network modeling approach to circuit optimization and statistical design, IEEE Trans. Microwave Theory Technol. 43 (6) (1995) 1349–1358.

Ignacio Santamaría was born in Vitoria, Spain, in 1967. He received the Telecommunication Engineer degree and the Doctor degree from the Universidad Politécnica de Madrid (UPM), Spain, in 1991 and 1995, respectively. In 1992 he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interests include digital signal processing, nonlinear systems and neural networks.



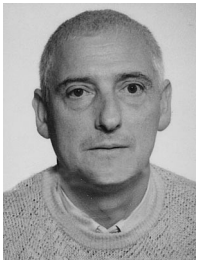
Marcelino Lázaro was born in Carriazo, Spain, in 1972. He received the Telecommunications Engineer degree from the Universidad de Cantabria, Spain, in 1996. Currently he is working at the Departamento de Ingeniería de Comunicaciones of the Universidad de Cantabria as a postgraduate student. His research interests include digital signal processing, nonlinear modeling and neural networks.

Carlos Pantaleón was born in Badajoz, Spain, in 1966. He received the Telecommunication Engineer degree and the Doctor degree from the Universidad Politécnica de Madrid (UPM), Spain, in 1990 and 1994, respectively. In 1990 he joined the Departamento de Ingeniería de Comunicaciones at the Universidad de Cantabria, Spain, where he is currently an Associate Professor. His research interests include digital signal processing, nonlinear systems and neural networks.



José Angel García (S'98) was born in Havana, Cuba, in 1966. He graduated (with honors) in Telecommunications Engineering from "José A.Echeverría" Institute of Technology (I.S.P.J.A.E.) in 1988. From 1988 to 1991 he was a Radio System Engineer at an HF Communication Center. In 1991, he was appointed Instructor Professor at I.S.P.J.A.E.

Since 1995, he has been a Ph.D. Student at University of Cantabria, Spain, with a MUTIS grant. His main research interests include nonlinear characterization and modeling of active devices as well as nonlinear circuits analysis tools for amplifiers and mixers applications.



Antonia Tazón Puente (M'92) was born in Santander, Spain, in 1951. He graduated in 1978 and received the Doctor of Physics degree in 1987, both from the University of Cantabria, Santander, Spain. From 1991 to 1995 he was professor at the Department of Electronics University of Cantabria) and since 1996 he is professor at the Department of Communication Engineering also of the University of Cantabria, Spain. In 1985, from March to October, in 1986, from April to July, he carried out stages at the IRCOM Department (University of Limoges, France) working in nonlinear modeling and load-pull techniques.

He has participated in Spanish and European projects in the nonlinear modeling (Esprit project 6050 MANPOWER) and microwave and millimeter wave communication circuits and systems (Spanish Project PlanSAT, European Project CABSINET, etc).

He has carried out research on analysis and optimization of nonlinear microwave active devices and circuits in both hybrid and monolithic technologies. Currently, his main research interests is the active microwave circuits, mainly in the area of linear and large-signal modeling and small signal intermodulation of GaAs and Si–Ge devices and their applications in nonlinear computer design.

Angel Mediavilla Sánchez (M'92) was born in Santander, Spain, in 1955. He graduated (with honors) in 1978 and received the Doctor of Physics degree in 1984, both from the University of Cantabria, Santander, Spain. From 1980 to 1983 he was Ingénieur Stagière at Thomson-CSF, Corbeville, France. Now, he is professor at the Department of Electronics of the University of Cantabria.

He has a wide experience in analysis and optimization of nonlinear microwave active devices and circuits in both hybrid and monolithic technologies. He has participated in Spanish and European projects in nonlinear modeling (Esprit project 6050 MANPOWER) and microwave and millimeter wave communication circuits and systems (Spanish Project PlanSAT, European Project CABSINET, etc).

His current research fields are on active microwave circuits, mainly in the area of nonlinear modeling of GaAs devices and their applications in large-signal computer design.