Mathematical & Computational Applications, Vol. 1, N0. 1, pp 113-118, 1996

NEURAL NETWORK APPROACH FOR THE CHARACTERISATION OF THE ACTIVE MICROWAVE DEVICES

Hamid TORPİ[1] Filiz GÜNEŞ[1] Fikret GÜRGEN[2]

[1]*Yıldız Technical University , Electronics &Communication Eng. Dept.,80670 MASLAK-ISTANBUL/TÜRKİYE*
[2]*Boğaziçi University, Computer Eng. Dept. ,BEBEK-ISTANBUL/TÜRKİYE*

Abstract

Small-signal and noise behaviour of an active microwave device is modeled through the neural network approach for multiple bias/configurations.Here ,the device is modelled by a black box whose small signal and noise parameters are evaluated through a neural network,based upon the fitting of both of these parameters for the multiple bias or configuration.The concurrent modelling procedure does not require to solve device physics equations repeatedly during optimization .Compared to the existing device modelling techniques,the proposed approach has the capability to make high-dimensional models for highly nonlinear devices.

I. INTRODUCTION

Aim of the work

The aim of this work is to model a microwave transistor by a black-box for which small-signal and noise parameters are evaluated through a neural network , based upon the fitting of both of these parameters to the corresponding measured data over the whole operational range from DC to more than 10 GHz for multiple bias of various types of configuration.So the stages of the work can be ordered as follows:

(i) Establish a novel neural network of feedforward type with a single hidden layer,

(ii) using back-propagation and nonlinear types of activation functions ,train the network for both the signal-noise behaviours over the operational bandwidth for multiple bias and multiple configuration of any type of microwave transistor.

(iii) Establish performance measure of the model,

(iv) Predict the small-signal and noise behaviours at any operation frequency around any bias condition of any type of configuration using the neural network which has already been trained to make functional approximations of the device nonlinear characteristics in the vicinities of the chosen bias points.

From the classical point of view ,unified small-signal-noise equivalent circuit for a microwave transistor can be divided in to two parts:Extrinsic circuits and intrinsic circuit.The intrinsic circuit characterize the active region under the gate (or base) whose parameters are functions of bias conditions and device technological parameters,whereas the extrinsic parameters depend,at least to a first approximation,only on the technological parameters.If an unified circuit for a MESFET is considered ,the most important extrinsic parameters are the gate,source and drain inductances due to the bond wires and the gate,source and drain resistances.The four main intrinsic parameters are the input capacitance C_{GS} ,the transconductance g_m ,the output conductance g_d and the feedback capacitance C_{GD}.In addition ,the electrical behaviour of the intrinsic device requires the introduction of two more parameters:The intrinsic resistance R_i which can be related to distrubuted nature of the RC input network,and the delay τ,which is introduced in the expression for the current generator and corresponds to the time needed for the carriers to travel under the gate.

Briefly due to the intrinsic and extrinsic device properties,both the signal and noise parameters are the functions of the bias conditions,frequency,configuration types.The way to approximate these functions in the literature so far is considered in the forthcoming subsection.

Review the literature

The problem of approximating measured device parameters or device response has been formulated as an optimization problem with respect to the equivalent circuit of a proposed model.The traditional approach in modeling is almost entirely directed at achieving the best possible match between measured and calculated parameters.This approach has serious shortcomings in two frequently encountered cases.The first case is when the equivalent circuit parameters are not unique with responses selected and the second is when the nonideal effects are not modelled adequately,the latter causing an imperfect match even if small measurement errors are allowed for .In both cases ,a family of solutions for circuit model parameters may exist which produce reasonable and similar matches between measured and calculated responses.Besides,published literature is concerned with the equivalent circuit for the single-bias which are only either small-signal models or the noise behaviour descriptions based on existing signal equivalent circuits that have nothing to do with the device noise characteristics .In [1] and [2] these two behaviours are combined in an unified classical circuit model for only a single-bias.A recent work [3] combines the signal and noise parameters in aneural network model over the fairly large operation band at a single bias point.

II. NEURAL NETWORK MODEL

Signal-Noise Behaviour of a Microwave Transistor

S and **N** parameter data measured at the multiple bias condititons (V_{DS},I_{DS}) for the configuration types (0.2,0.5,0.8) is all used to train the neural network.The amount of data used in the training and iteration number are altogether optimised against the error The measured **S** and **N** parameter data around a bias point for a type of configuration can be arranged in a table-form function as follows:

$$\begin{bmatrix} f_1 & : & S^{(1)} & N^{(1)} \\ f_2 & : & S^{(2)} & N^{(2)} \\ & & & \\ f_N & : & S^{(N)} & N^{(N)} \end{bmatrix} \quad (1)$$

where $S^{(1)}$, $N^{(1)}$; ... ;$S^{(N)}$,$N^{(N)}$ are respectively, the scattering and noise vectors at the f_1, ,...,f_N measurement frequencies,and $S^{(N)}$ and $N^{(N)}$ performance vectors can be given as follows:

$$\left[S^{N}\right]^{t} = \left[|S_{11}|^{(N)} \ \varphi_{11}^{(N)} \ |S_{12}|^{(N)} \ \varphi_{12}^{(N)} \ |S_{21}|^{(N)} \ \varphi_{21}^{(N)} \ |S_{22}|^{(N)} \ \varphi_{22}^{(N)}\right]$$

$$\left[N^{(N)}\right]^{t} = \left[F_{opt}^{(N)} \quad |\Gamma_{opt}|^{(N)} \quad \varphi_{opt}^{(N)} \quad R_{N}^{(N)}\right] \quad (2)$$

After having completed the training process,the performance vectors $S^{(K)}$, $N^{(K)}$ at a desired frequency f_k at the conditions (V_{CE} ,I_C ,CT or V_{DS},I_{DS} ,CT) for any configuration type among the trained ones can be obtained from the network output by inputing the frequency f_k bias configuration type which is defined by the numbers (0.2,0.5,0.8).If $S^{(K)}$, $N^{(K)}$ are unmeasured they are determined by the generalization process of the neural network,which can be considered as the ability of the network to give good outputs to inputs it has not been trained on.In our application ,the signal-noise neural network can generalize the performance not only at a single operation frequency of the trained bias condition ,at the same time in the whole operation band of the untrained bias condition.The first may be named as the single frequency genaralization ,while the latter is called whole band generalization ,worked examples of which will be given in the result section.

The multi-bias and configuration signal-noise neural network

We use a novel neural network of feedforward type with a single hidden layer having the same number of nodes as the output layer.Let n,N_h and N_o be respectively the number of nodes in the input ,hidden and output layers.In the signal-noise neural network n=4 with the frequency,bias condition and the type of configuration CT,N_h=N_o=12 which are the signal and noise vectors given by (2). (Fig.2)

The signal resulted from the hidden layer to the i th output node can be expressed inthe form of

$$\Phi_i\left(\mathbf{x},\mathbf{T}_i,\mathbf{W},\theta_h\right) = \sum_{h=1}^{N_h} T_{hi} g_h\left(\mathbf{x},\mathbf{W}_h,\theta_h\right) \quad (3)$$

and the net output of the i th output node is obtained as follows

$$\phi_i\left(\mathbf{x},\mathbf{T}_i,\mathbf{W},\Theta\right) = V_i f_i\left(\Phi_i + \Theta_i\right) \quad (4)$$

where g_h and f_i are the basis functions for the h th hidden node and the i th output node , respectively , which are sigmoid type of nonlinear functions in our case,e.g. $g_h(\mathbf{W}_h,\mathbf{x})$ can be expressed in the following form:

$$g_h\left(\mathbf{x},\mathbf{W}_h,\theta_h\right) = \frac{1}{1+\exp\left(-\sum_{i=1}^{n} x_i W_{ih} - \theta_h\right)} \quad (5)$$

In equations 3-5,**x** is the input vector of n dimensions:

$$\mathbf{x}=[x_1,x_2,\ldots,x_n]^t \quad (6)$$

$\mathbf{T}_i$ is the weighting vector between the i th output node and the hidden layer:

$$T_i = \left[T_{1i}\ T_{2i}\ T_{3i}\ \ T_{hi}\ \ T_{N_h i}\right]^t \quad (7)$$

W is the weighting matrix between the hidden and input layer:

$$\mathbf{W}=[\mathbf{W}_1,\mathbf{W}_2\ \ldots\ \mathbf{W}_h\ \ldots\ \mathbf{W}_{Nh}\,]^t \quad (8)$$

where $\mathbf{W}_h$ vector is the weight(ing) vector between the input layer and the h th hidden node and can be given by:

$$\mathbf{W}_h=[W_{1h},W_{2h},\ \ldots\ W_{nh}]^t \quad (9b)$$

θ_h ,Θ_i are the local memories belonging to the h th hidden and i.th output nodes,respectively.In the eqn. (4) **V** is the weighting factor of the output layer:

$$\mathbf{V}=[V_1,V_2,\ldots,V_{No}]^t \quad (9b)$$

Determination of the network parameter matrix ***P***

If parameters of the network architecture is denoted **P** ,the network parameter matrix **P** will have $N_h$xn +N_hxN_o+N_h+N_o elements which consist of weighting factors between the input and hidden layers and the hidden and output layers ,the local memories of the hidden and output nodes.The training process can be defined as computation of the network parameter matrix **P** so that the error function which is

$$E(\mathbf{P}) = \sum_{k=1}^{N_S} E^{(k)} = \sum_{k=1}^{N_S}\left[\frac{1}{2}\sum_{j=1}^{N_O}\left(y_j^{(k)} - \phi_j^{(k)}\right)^2\right] \quad (10)$$

is minimised ,where $y_j^{(k)}$ and $\phi_j^{(k)}$ are respectively ,the measured and predicted values of the j th output at the training frequency f_k.This training process is also called backpropagation and it is an 'on-line' process whose update equations for T_{hj},W_{ih},θ_h can be given as follows:

$$T_{hj}^{(k+1)} = T_{hj}^{(k)} - \eta\frac{\partial E^{(k)}}{\partial T_{hj}} + \alpha\left(T_{hj}^{(k)} - T_{hj}^{(k-1)}\right) \quad (11.1)$$

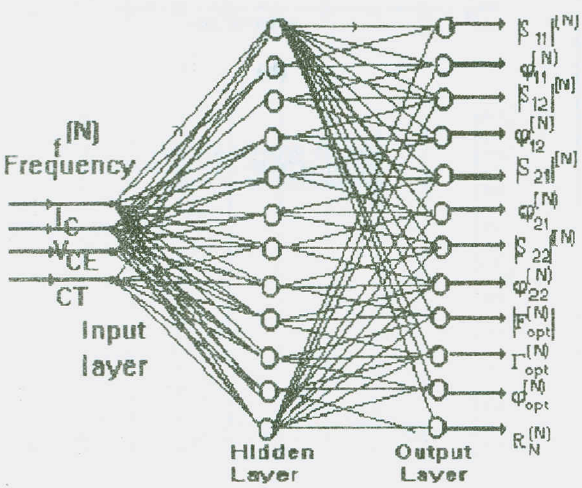
$$W_{ih}^{(k+1)} = W_{ih}^{(k)} - \eta \frac{\partial E^{(k)}}{\partial W_{ih}} + \alpha\left(W_{ih}^{(k)} - W_{ih}^{(k-1)}\right) \quad (11.2)$$

$$\theta_h^{(k+1)} = \theta_h^{(k)} - \eta \frac{\partial E^{(k)}}{\partial \theta_h} + \alpha\left(\theta_h^{(k)} - \theta_h^{(k-1)}\right) \quad (11.3)$$

and the similiar equations can be written for Θ_i and V_i.In (11.1)-(11.3) η and α are positive-valued learning rate and momentum,respectively.Thus we start any set the network parameters and then repeatedly change each parameter by an amount proportional to the terms $\frac{\partial E^{(k)}}{\partial T_{hj}}$, $\frac{\partial E^{(k)}}{\partial W_{ih}}$, $\frac{\partial E^{(k)}}{\partial \theta_{ih}}$ according to the eqns. (11.1-11.3) and assume that the training is completed when the error fails to decrease any further.In this case we take the best so far.

Fig .1 Multi-bias and configuration signal-noise neural network

The sensitivity through the neural network with respect to T_{hj} , W_{ih},θ_h can be given as follows



$$\frac{\partial E^{(k)}}{\partial T_{hj}} = \frac{\partial\left[\frac{1}{2}\sum_{j=1}^{N_O}\left(y_j^{(k)} - \phi_j^{(k)}\right)^2\right]}{\partial T_{hj}} = -V_j\left(y_j^{(k)} - \phi_j^{(k)}\right)f_j\left(1-f_j\right)g_h$$

$$\frac{\partial E^{(k)}}{\partial T_{hj}} = \delta_j^{(3)} g_h \quad (12.1)$$

$$\frac{\partial E^{(k)}}{\partial W_{ih}} = \sum_{j=1}^{N_O} \frac{\partial E^{(k)}}{\partial \phi_j}\frac{d\phi_j}{df_j}\frac{df_j}{d\Phi_j}\frac{\partial \Phi_j}{\partial W_{ih}} = \sum_{j=1}^{N_O} -\delta_j^{(3)} \frac{\partial \Phi_j}{\partial W_{ih}}$$

$$\frac{\partial E^{(k)}}{\partial W_{ih}} = g_h\left(1-g_h\right)\sum_{j=1}^{N_O} -\delta_j^{(3)} T_{hj} x_{ki} = \delta_h^{(2)} x_{ki} \quad (12.2)$$

$$\frac{\partial E^{(k)}}{\partial \theta_h} = \sum_{j=1}^{N_O} \frac{\partial E^{(k)}}{\partial \phi_j}\frac{d\phi_j}{df_j}\frac{df_j}{d\Phi_j}\frac{\partial \Phi_j}{\partial \theta_h}$$

$$\frac{\partial E^{(k)}}{\partial \theta_h} = g_h\left(1-g_h\right)\sum_{j=1}^{N_O} -\delta_j^{(3)} T_{hj} = \delta_h^{(2)} \quad (12.3)$$

$$\frac{\partial E^{(k)}}{\partial V_j} = \frac{\partial E^{(k)}}{\partial \phi_j^{(k)}}\frac{\partial \phi_j^{(k)}}{\partial V_j} = -\left(y_j^{(k)} - \phi_j^{(k)}\right)f_j\left(\Phi_j + \Theta_j\right)$$

$$\frac{\partial E^{(k)}}{\partial V_j} = \delta_j^{(3)}\left(1-f_j\right)^{-1} \quad (12.4)$$

$$\frac{\partial E^{(k)}}{\partial \Theta_j} = -\left(y_j^{(k)} - \phi_j^{(k)}\right)f_j\left(1-f_j\right) = \delta_j^{(3)} V_j^{-1} \quad (12.5)$$

where $\delta_h^{(2)}$ and $\delta_j^{(3)}$ represent local gradients at individual node in the second and third layer ,respectively.

III.PERFORMANCE MEASURE AND RESULTS

To evaluate the quality of the fit to measured data the following error terms are found to be convenient :

$$E_{S_{ij}} = \frac{1}{n}\sum_{k=1}^{n} \frac{\left|S_{ij_{meas}}^k - S_{ij_{meas}}^k\right|}{\left|S_{ij_{meas}}^k\right|} \quad (13)$$

$$E_{N_i} = \frac{1}{n}\sum_{i=1}^{n} \frac{\left|N_{i_{meas}}^k - N_{i_{predict.}}^k\right|}{\left|N_{i_{meas.}}^k\right|} \quad (14)$$

Where S_{ij} and N_i are ,respectively the signal and noise parameters , and n is the number of discrete frequencies used .Total average error can be defined as the average of the signal and noise errors:

$$E_T = \frac{1}{4}\sum_{i=1}^{4} E_{i_{signal}} + \frac{1}{3}\sum_{i=1}^{3} E_{i_{noise}} \quad (15)$$

Distribution of errors with frequency for the whole band generalization at multiple bias points for the common collector configuration is given in Fig.3.Simulation results of NE02135 (iter.num.400000) Transistor are given in Fig.4 a-d which shows Distribution of errors with frequency for the whole band generalization at multiple bias points for the com-mon emitter configuration. In addition simulation results of NE219 transistor are given in Fig.5 a-f which shows variations of S parameters with frequency from 2 -6 GHz for the V_{CE}=8V and I_C=10,20 and 30 mA at the common emitter configuration which show quite good agree-ment of the signal parameters over the operation band-width. The graphs include variations of S parameters and noise parameters with frequency from 0.5 -4 GHz for the V_{CE}=10V and I_C= 5, 10, 20 and 30 mA at the common emitter

10, 20 and 30 mA at the common emitter configuration.Finally variations of S parameters with respect to bias point is given for various constant frequencies.(Fig.6)

Acknowledgement: This work was supported by the Yıldız Technical University Research Fund. Project number: 92-A-04-03-12.

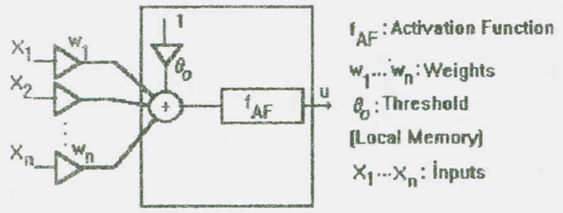


Fig. 2 A perceptron node

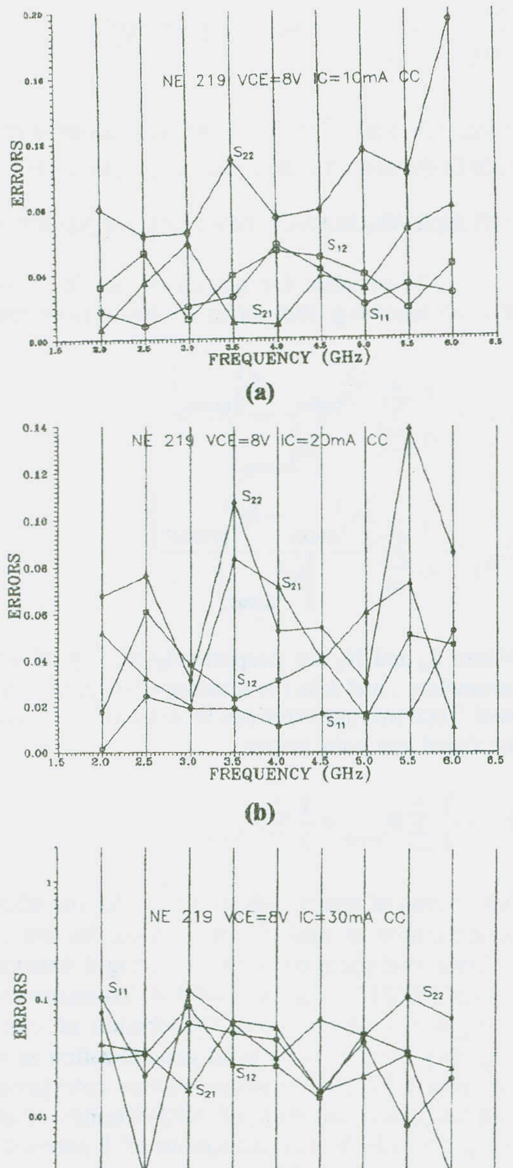


Fig. 3 Distribution of errors with frequency for the whole band generalization at multiple bias points for the common collector configuration (NE219)

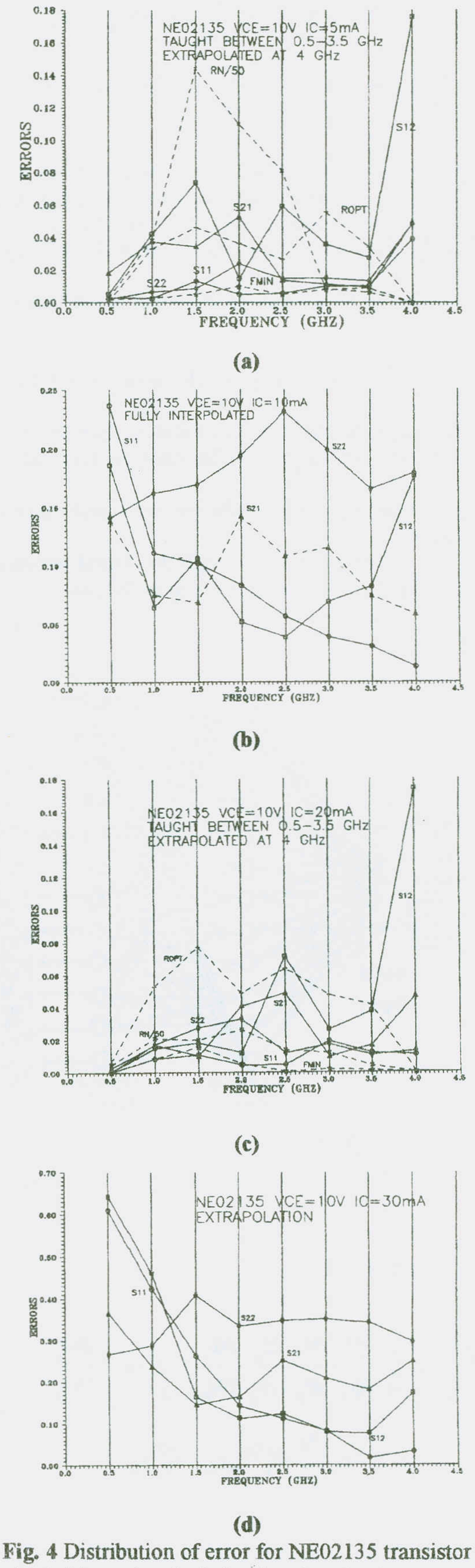
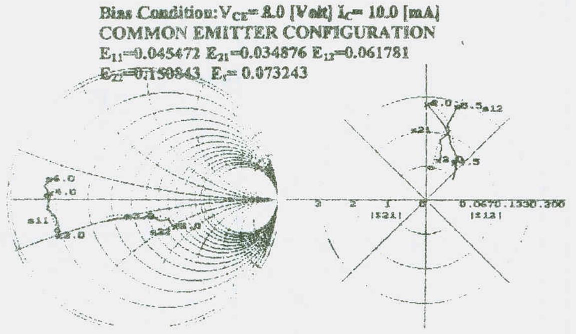
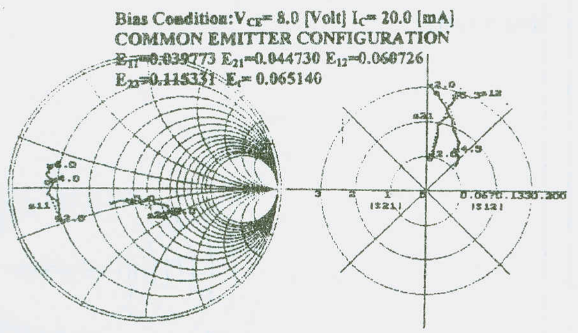


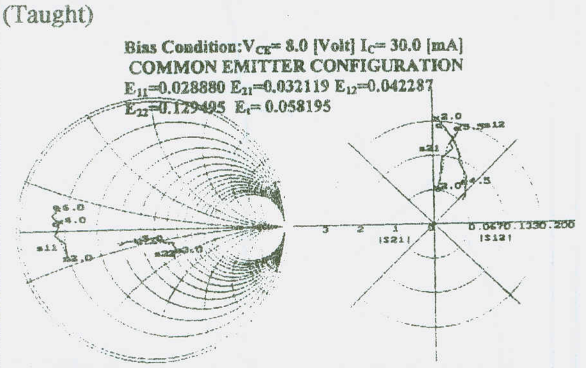
Fig. 4 Distribution of error for NE02135 transistor at multiple bias points both for signal and noise parameters (if measured noise data exist)



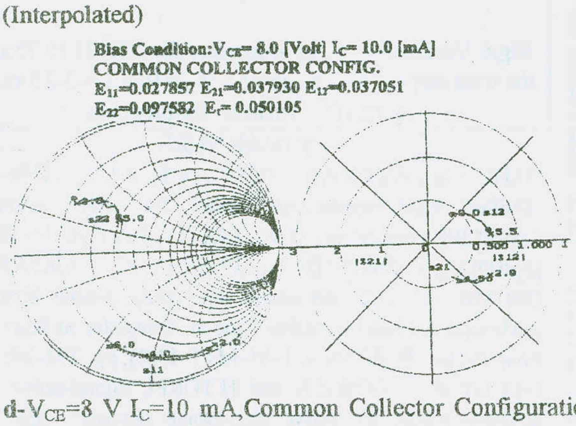
a- V_{CE}=8 V I_C=10 mA, Common Emitter Configuration (Taught)



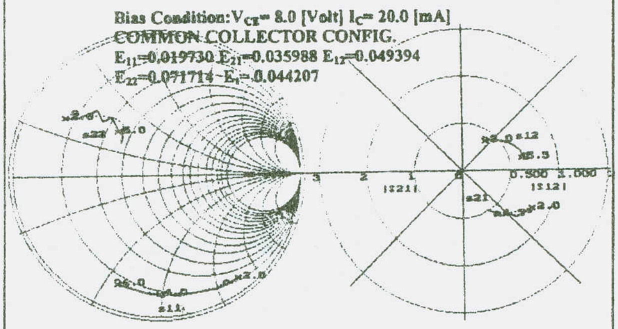
b- V_{CE}=8 V I_C=20 mA, Common Emitter Configuration (Taught)

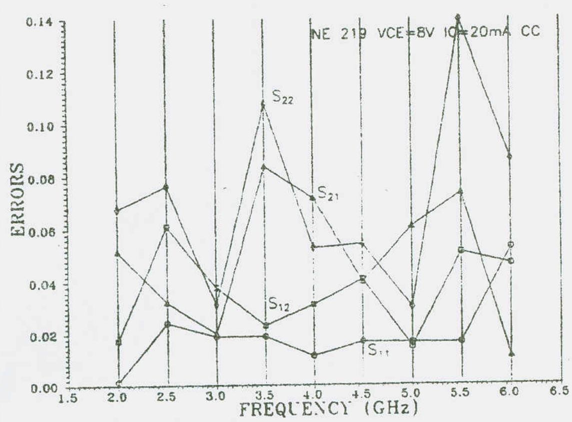


c-V_{CE}=8V I_C=30 mA, Common Emitter Configuration (Interpolated)

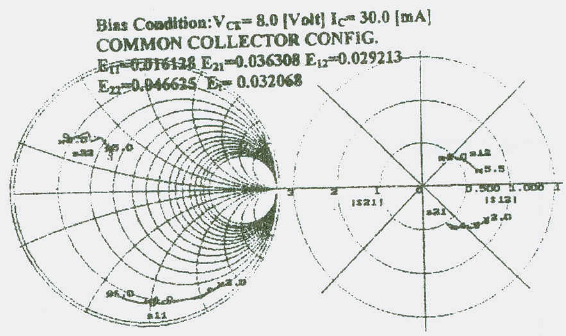


d-V_{CE}=8 V I_C=10 mA,Common Collector Configuration (Interpolated)



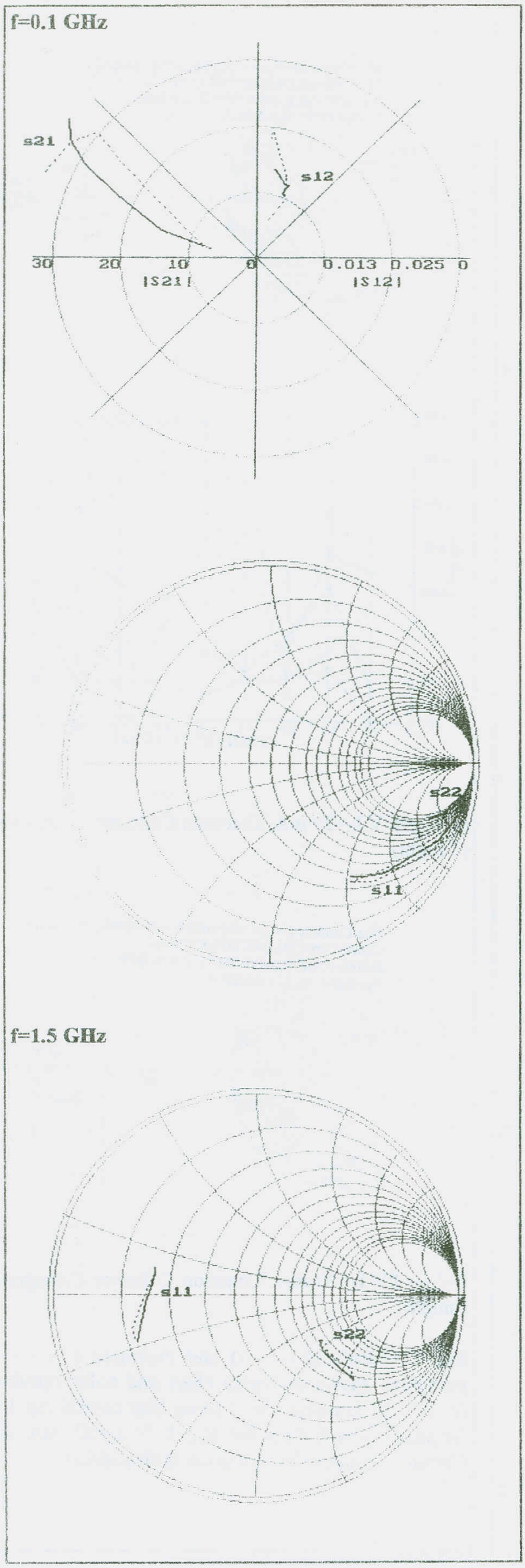


e- V_{CE}=8 V I_C=20 mA, Common Collector Configuration (Taught)



f-V_{CE}= 8V I_C=30 mA, Common Collector Configuration (taught)

Fig. 5 Calculated (-------) and measured (- - - - -) S parameter shown on Smith chart and polar coordinates for NE219 Transistor at various bias conditions .Error-frequency distribution for V_{CE}=8 V I_C=20 mA at the Common Emitter Configuration is also added.



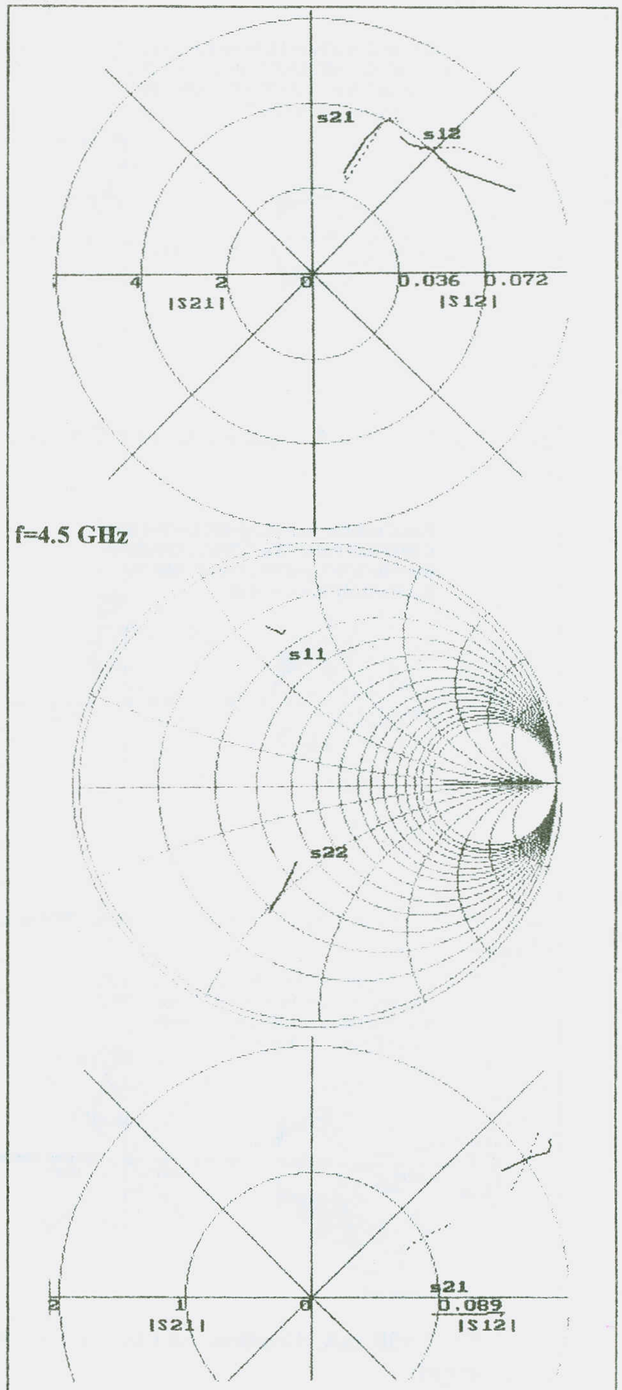


Fig.6 Variations of S parameters for NE02135 Transistor with respect to bias point (V_{CE}=10V I_C=2-25 mA)is given for various frequencies.

REFERENCES

[1].HU Z.R., YANG Z.M., FUSCO V.F., J.A.C. STEWART , "Unified small-signal-noise model for active microwave device",IEE Proceedings-G,Vol.140, (1) , 1993 ,pp. 55-60.

[2].ROUX,J.P.,ESCOTTE L.,PLANA R., GRAFFEUIL, DELAGE S.L , "Small-Signal and Noise Model Extraction Technique for heterojunction Bipolar Transistor at Microwave Frequencies",IEEE Trans. 1995-MTT-43(2), pp. 293-298.

[3].F.GÜNEŞ,F.GÜRGEN and H.TORPİ,"Signal-noise neural network model for active microwave devices", IEE Proc.-Circuits Devices and Syst.,Vol.143,No.1,,February 1996, pp. 1-8.