[35] B. A. Pearlmutter, "Gradient Calculations for Dynamic Recurrent Neural Networks: A Survey," *IEEE Trans. on Neural Networks*, Vol. NN-6, 1995, pp. 1212–1228.

[36] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme Recognition Using Time Delay Neural Networks," *IEEE Trans. on Acoustics, Speech, Signal Processing*, Vol. ASSP-37, 1989, pp. 328–339.

[37] L. Leistritz, M. Galicki, H. Witte, and E. Kochs, "Training Trajectories by Continuous Recurrent Multilayer Networks," *IEEE Trans. on Neural Networks*, Vol. NN-13, no. 2, 2002, pp. 283–291.

[38] G. Casinovi, and A. Sangiovanni-Vincentelli, "A Macromodeling Algorithm for Analog Circuits," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-10, 1991, pp. 150–160.

[39] P. K. Gunupudi, and M. S. Nakhla, "Model-Reduction of Nonlinear Circuits using Krylov-Space Techniques," *Proc. IEEE Int. Design Automation Conf.*, New Orleans, LA, June 1999, pp. 13–16.

[40] J. Vlach, and K. Singhal, *Computer Methods for Circuit Analysis and Design*, New York: Van Nostrand Reinhold, 1993.

[41] *NeuroModeler Version 1.02*, Prof. Q. J. Zhang, Department of Electronics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada.

[42] *ADS-Advanced Design System* Version 2002, Agilent Technologies, Santa Rosa, CA, 2002.

[43] K.S. Kundert, G.B. Sorkin, and A. Sangiovanni-Vincentelli, "Applying Harmonic Balance to Almost-Periodic Circuits," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-36, 1988, pp. 366–378.

[44] M. Deo, J. J. Xu, and Q. J. Zhang, "A New Formulation of Dynamic Neural Network for Modeling of Nonlinear RF/Microwave Circuits," *Proc. European Microwave Conf.*, Munich, Germany, Oct. 2003, pp. 1019–1022.

[45] M. C. E. Yagoub, and H. Baudrand, "Optimum Design of Nonlinear Microwave Circuits," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-42, 1994, pp. 779–786.

[46] *NeuroADS*, Prof. Q. J. Zhang, Department of Electronics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada.

Chapter 7

Behavioral Modeling from the Perspective of Nonlinear Dynamics

John Wood and David E. Root
Worldwide Process and Technology Centers,
Agilent Technologies, Inc., Santa Rosa, California

Nicholas B. Tufillaro
Agilent Laboratories,
Agilent Technologies, Inc., Palo Alto, California

7.1 INTRODUCTION

Over the past two decades, advances in nonlinear simulation techniques have enabled the efficient numerical solution of problems important in the design of new microwave and RF integrated circuits. Some key advances have been the development of modern harmonic balance simulators, and more recently, of transient envelope simulators. Harmonic balance [1] allows the efficient simulation of large-signal steady-state circuits in the frequency domain, achieving great efficiency gains over traditional time-domain simulators like SPICE for steady-state problems with large numbers of frequency components. Such multisine problems are often found in the design of microwave circuits. The recent availability of transient envelope simulators [2, 3] has allowed the efficient simulation of problems for which the typical spectra can be represented by a set of several discrete tones and time-dependent modulation around them. This modulation is a common characteristic of modern communication circuits.

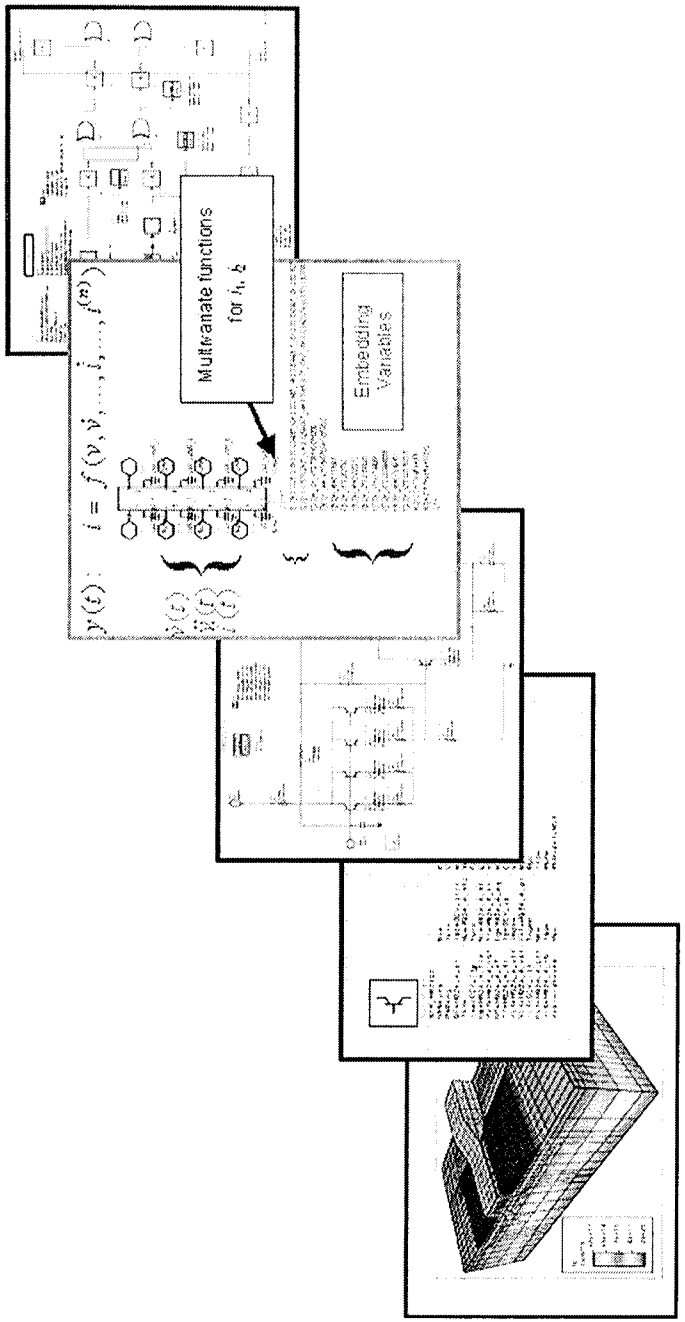


Figure 7.1 A modeling hierarchy, starting at the bottom with a device model described by the detailed semiconductor physics; this is abstracted to a circuit-level transistor model that describes the terminal behavior through equivalent-circuit or phenomenological equations. The transistor model is used to design circuits and ICs efficiently in circuit simulators, but is too complex to use in the system level simulation at the top of the hierarchy. To bridge circuit and system simulation environments, we use a behavioral model of the circuit.

Even so, modern microwave and wireless communications systems are nowadays too complex to permit the complete simulation of the nonlinear behavior at the transistor level of description. The combination of a large number of nonlinear devices — the transistors in the circuits — and a complex signal modulation mean that the number of nonlinear equations that must be solved at each iteration step in the simulation is prohibitively large. The memory requirements alone discourage all but the most determined (or desperate), and even then the combination of the multiple nonlinearities in the whole system will generally preclude the simulation from reaching convergence. And it may take several hours for the simulation finally to crash. This problem presents a significant productivity bottleneck for system and subsystem design engineers. A solution to this design productivity bottleneck is to design at a higher level of abstraction at each level in the system hierarchy.

A typical system design and modeling hierarchy is depicted in Figure 7.1. At the bottom is the device, and at the top is the complicated module or subsystem. A "top-down" design methodology propagates specifications down the hierarchy. Conversely, "bottom-up" verification is the process of validating overall system performance based on the performance of lower-level components and their configuration.

At the bridge between the three-dimensional (3-D) physical device simulation and the transistor circuit is the circuit model for the transistor device. This model is often devised using simplified physical equations, or phenomenological equations that describe the terminal behavior of the transistor in terms of physical parameters or observables, such as the dimensions of the device and the charge carrier (electron) properties such as mobility. This simplified physical description for the transistor's electrical behavior enables the simulation of fairly complex circuits within reasonable time and computer resources that would be impossible using the 3-D physical model, which is based on partial differential equations that require involved and time-consuming solution techniques.

Similarly, at the bridge between the transistor circuit and the multichip module or RFIC is the behavioral model, which is used to describe the nonlinear circuit blocks or ICs in the system. The behavioral models are simplified models of the essential nonlinear behavior of the complex subcircuits; this simplification means that these models will execute more quickly, and use much less memory than if an entire complex subsystem was simulated at the transistor level. The critical need for nonlinear modeling techniques is a recent development driven by the increased size

and complexity of ICs in the RF regime, as well as the adoption of more complex signal modulation techniques. The availability of such nonlinear modeling techniques will enable designers to make use of the advances in the simulation technology at higher levels of the design hierarchy.

In this chapter we describe a new and systematic methodology for generating nonlinear behavioral models that is based on observing the nonlinear dynamics of the device or IC. Our approach to what is a fundamentally nonlinear discipline has a sound theoretical background in system identification and computational geometry. These behavioral models arise from a black-box approach, where we are concerned only with describing the dynamical behavior of the circuit that is observed at its accessible terminals. This is in contrast to the traditional white-box approach, where detailed knowledge of the device physics or circuit configuration and operation is used to minimize the number of equations that describe the essential properties of the circuit or device. Indeed, we advocate this black-box approach even when such details are known, as it is based on the observable dynamics of the system, which are generally of much lower order than the number of degrees of freedom as may be inferred from the physics of the device.

Our approach provides a systematic framework for creating black-box behavioral models that can be applied in a general way; it is not an *ad hoc* black-box modeling approach. This methodology also provides a framework for understanding other modeling approaches such as the application of dynamic neural networks (DNNs) to the behavioral modeling of RFICs, as described in Chapter 6.

The pervading themes of our methodology include the notion of describing the relationships between signals or waveforms, encompassed by functional analysis; the reconstruction of the observable nonlinear dynamics of the device from the measured data, and the choice of the excitations required to enable the observation of these dynamical behaviors.

We will illustrate our methodology by creating and validating a model of a real microwave IC amplifier. The modeling procedures that we outline are very general: the test signal design, analysis, model generation, and simulator implementations are generic and can be applied to amplifiers, mixers, modulators, and other microwave components or subsystems. The resulting models are:

(1) "transportable" [4] — in other words, usable in a range of system and simulation environments, and not restricted to a limited domain of applicability;
(2) "cascadable" [5] — so that the cascade of two (or more) behavioral models performs faithfully with respect to the performance of the cascade of the respective transistor-level circuits.

Since cascading nonlinear components can create a wide variety of environments for the individual behavioral models, "cascade-ability" implies a certain degree of transportability.

7.2 NONLINEAR DYNAMICS

The origin of dynamical nonlinearities can be seen using a relatively simple example from semiconductor physics: the field-effect transistor (FET). The FET device can be described in three dimensions (see Figure 7.1, for example) using fundamental physical concepts including the quantum theory of matter, electromagnetism, and thermodynamics, resulting in a set of coupled partial differential equations that govern the electrical behavior of the device under consideration. By using phenomenological approximations such as field-independent mobility, the gradual-channel approximation, and utilizing the symmetry of the structure, the constitutive equations that describe the currents and charges (or capacitances) can be derived [6]:

$$I_d^{DC}\left(V_{gs},V_{ds}\right)=\frac{W\mu qN_D a}{\varepsilon L}\left[V_{ds}-\frac{2}{3}\sqrt{\frac{2\varepsilon}{qN_D a^2}}\left\{\left(V_{ds}+\varphi-V_{gs}\right)^{3/2}-\left(\varphi-V_{gs}\right)^{3/2}\right\}\right] \quad (7.1)$$

$$C(V)=WL\sqrt{\frac{q\varepsilon N_D}{2(\varphi-V)}}$$

From these relations a simple intrinsic model of the FET can be derived, shown in Figure 7.2, from which the terminal equations reveal the following dynamical relations:

$$\begin{aligned} I_d(t) &= I_d^{DC}\left(V_{gs}(t),V_{ds}(t)\right)-C\left(V_{gd}(t)\right)\frac{dV_{gd}(t)}{dt} \\ I_g(t) &= C\left(V_{gs}(t)\right)\frac{dV_{gs}(t)}{dt}+C\left(V_{gd}(t)\right)\frac{dV_{gd}(t)}{dt} \end{aligned} \quad (7.2)$$

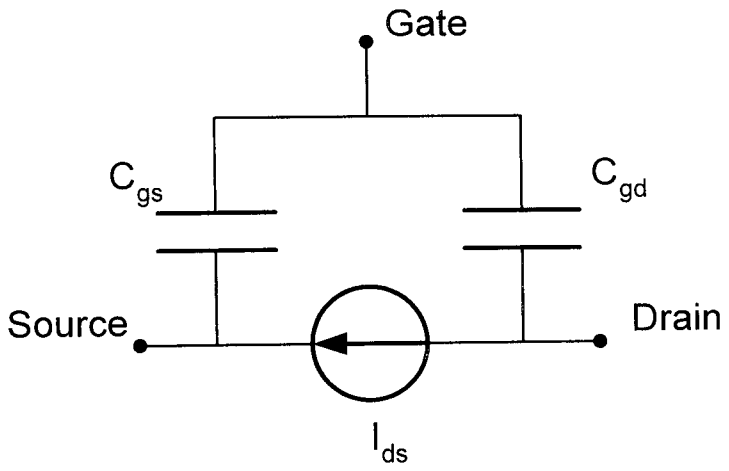


Figure 7.2 Simple intrinsic FET model schematic.

The capacitance expressions yield time-dependent currents that depend on the time rate of change of the applied voltage signals — these describe the dynamic nature of the terminal behavior. The current is dependent not only on the applied voltage, but also its time derivative: the current is dependent on the shape of the voltage waveform,

$$I(t) = f\left(v(t), \frac{dv(t)}{dt}\right) \tag{7.3}$$

7.2.1 Functional Analysis

Fundamental to our approach is the notion that the behavioral model relates waveforms to waveforms. That is, the output signal is not an instantaneous function only of the input signal. Rather, the output depends upon the shape of the input waveform, or equivalently, the output depends on the value of the input and past values of the input, and even past values of the output itself: the concept of *memory*. We write this in *functional* notation according to (7.4). Here we assume the voltage signal is the input and the current is the output:

$$I(t) = F[V(t)] \quad \text{or, as an implicit functional,} \quad F\left[V(t), I(t)\right] = 0 \tag{7.4}$$

How we write the details of this functional depends on the nature of the system or component being modeled. Take, for example, a nonlinear resistor: the current is given by the instantaneous value of the applied voltage. The details, or the shape, of the voltage signal are unimportant, there are no dynamics in this system. In this case, the current-voltage relationship can be expressed as a simple algebraic function:

$$I(t) = f\left(V(t)\right) \tag{7.5}$$

If we now place a capacitor in parallel with the resistor, the current flow depends on both the instantaneous voltage (across the resistor) and also the time derivative of the voltage (across the capacitor). The detailed shape of the voltage signal is now important. The current is now expressed as a functional of the voltage signal, as given by (7.6):

$$I(t) = F[V(t)] = f(V(t), \dot{V}(t)) \tag{7.6}$$

This is an example of what is termed a *static* functional: the output depends upon the input signal only, which is known as a static relationship. But the output depends on the shape of the voltage signal, as expressed through the time derivative of the voltage, and therefore this functional describes the dynamics present in the system: a dynamical model.

Now consider a series connection of the resistor and capacitor (either or both of which may be nonlinear components). It can be shown fairly easily that the current through the RC network depends on the time derivative of the applied voltage and the time derivative of the current itself:

$$I(t) = f\left(\dot{V}(t), \dot{I}(t)\right) \tag{7.7}$$

This is a more complicated class of models, known as a *feedback* model, since an output quality is required in the solution of the output itself. Feedback models depend on an internal state of the system, and also embody the notion of memory. We can write this as an implicit relationship — a functional of the current and voltage signals.

$$F[V(t), I(t)] = 0 \tag{7.8}$$

Again, this is a dynamical model as the functional relationship requires the output $I(t)$ to be found in terms of time derivatives of the input and output signals.

7.2.2 Dynamical Effects and Implications

By properly including the dynamical behavior in the model, the description of non-quasi-static effects can be handled accurately, and the limitations of quasi-static nonlinear models can be overcome. Examples of such dynamical behaviors include: electrothermal interactions in semiconductor devices, where the temperature of the device affects the RF gain; trapping effects in semiconductors; and bias circuit effects, where the impedance that is presented to the device varies with frequency. This latter effect may also include other frequency-dependent impedance variations, such as different termination impedances at various harmonics of the RF fundamental, for improving power amplifier efficiency, for example. The above effects relate the device behaviors at very different frequencies, or time-scales, and require sophisticated modeling techniques to describe fully. Simple quasi-static models cannot describe these phenomena.

The impact of such dynamical effects can be illustrated using an example of a resistor with a temperature-dependent resistance; the *I-V* relation is simply:

$$I(t) = V(t)/R(T) \tag{7.9}$$

and clearly, the resistor is electrically constant at a fixed temperature, T, yielding a linear *I-V* relationship.

The thermal constitutive relation for the resistance is given by

$$R(T) = R_0 + \Delta \cdot \left(T - T_0\right) \tag{7.10}$$

indicating that the resistance increases linearly with increasing temperature.

The dynamical equation that relates the temperature of the resistance to the input (dissipated) power, the electrothermal coupling equation, is written as the following nonlinear differential equation:

$$\tau \dot{T}(t) + \left(T(t) - T_0\right) = R_{th} \cdot \left(I(t)V(t)\right) \tag{7.11}$$

where R_{th} is the thermal resistance, and the $V.I$ product is the instantaneous power dissipated in the resistor. The time constant τ is a characteristic time of the resistor, describing how quickly the resistor can react to changes in the input power. While these equations yield a nonlinear ODE for $T(t)$, this can be solved exactly for simple cases, providing insight into the dynamical behavior. For example, solutions of (7.11) for a single-tone input voltage yield the following results:

(1) For input signal frequency $f \gg 1/\tau$, the output $I(t)$ is linear in $V(t)$, because the resistor temperature cannot change at the signal frequency.

(2) For input signal frequency $f \ll 1/\tau$, the output $I(t)$ is a square wave, that is, nonlinear with $V(t)$. The voltage is changing so slowly that the resistor temperature, and hence resistance, can track the voltage waveform exactly.

These waveforms are illustrated in Figure 7.3. This simple example illustrates that the nonlinear system can look linear under certain excitations, and so the excitation signals must be rich enough to stimulate the relevant dynamics that will be encountered in conditions of actual use, and enable accurate determination of the system dynamics. In this example, a single-tone high-frequency signal is unable to provide information about the slow thermal dynamics. While this may seem reasonable for a system that will only be used at high frequencies, if the system is excited with a two-tone signal with a tone separation of the order $1/\tau$, then the intermodulation product $f_1 - f_2$ will heat up the resistor and cause the resistance to change, thereby affecting the signals at the high frequencies. This coupling of the slow and fast dynamics can only be realized with a full nonlinear dynamical model.

A further example of the importance of including fully the dynamical behavior can be found in nonlinear modeling of FETs. In particular, it is noted that the small-signal RF transconductance g_m and output conductance g_o are not equal to the derivatives of the DC I-V relations for the device. Both g_m and g_o exhibit dispersion with frequency, and so the quasi-static model is incorrect. The implication of this is that the intermodulation and distortion products are not simply related to the higher-order derivatives of the static constitutive relations, but depend upon the time- (frequency-) dependent dynamics as well.

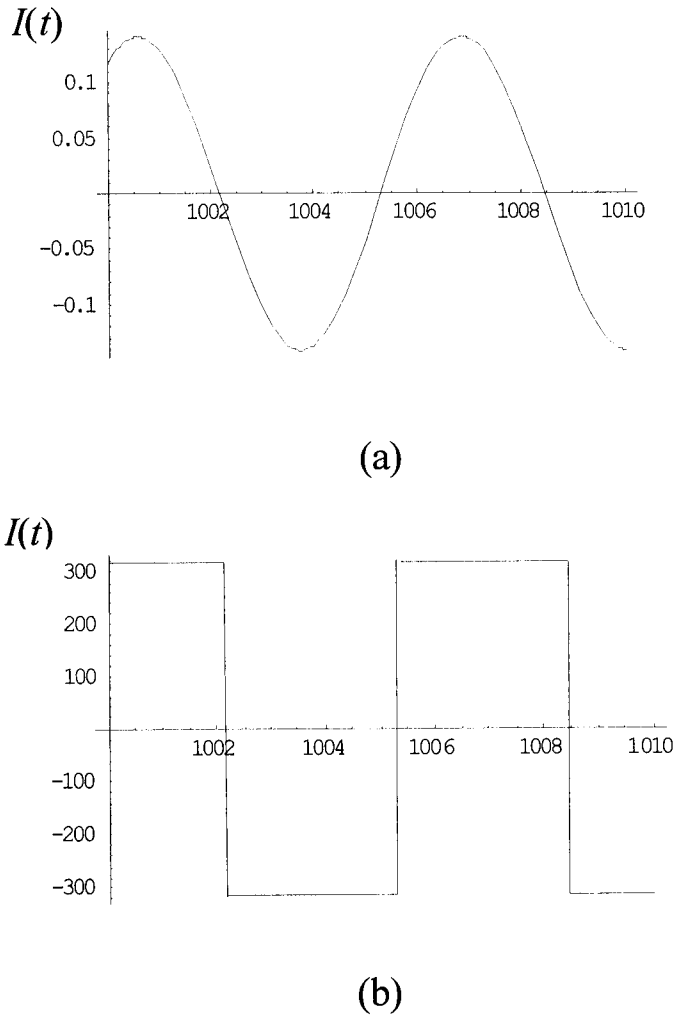


Figure 7.3 Current waveforms through the temperature-dependent resistor, for a single-tone sinusoidal drive voltage. (a) If the drive frequency is much greater than $(1/\tau)$, the current follows the voltage and the resistor appears linear. (b) If the drive frequency is much smaller than $(1/\tau)$, the resistor temperature and hence resistance can track the voltage, and the current is a square wave — nonlinear at this time- or frequency-scale. τ is the thermal time constant of the resistance.

7.3 BEHAVIORAL MODELING METHODOLOGY

The motivation for our approach to nonlinear systems identification goes under the rubric of nonlinear time series analysis (NLTSA) [7]. The suggestion to use this approach for describing input-output systems is due to Casdagli [8]. The key idea is to *embed* the measured or simulated stimulus and response variables in a higher dimensional space built not only from the measured data but also transforms of the measured data, in our case, their time derivatives that describe the local history of the signal. Due to a theorem of Takens, extended to the driven case by Stark [9], these embedded models can be faithful to the dynamics of the original system. In particular, deterministic prediction is possible from an embedded model that will mimic the dynamics of the actual system.

The models are formulated as implicit nonlinear ordinary time-differential equations, which are easily implemented in commercial microwave simulators, in the embedded variables:

$$f\left(i(t),\dot{i}(t),\ddot{i}(t),...,v(t),\dot{v}(t),\ddot{v}(t),...\right)=0 \tag{7.12}$$

The goal of the modeling process is thus to determine the significant embedding variables of the function, f, and then to find an efficient basis for the function approximation.

7.3.1 State Space Analysis

We can write the dynamical description of a nonlinear system in terms of a set of nonlinear ordinary differential equations.

$$\begin{aligned}\dot{\vec{x}}(t)&=\vec{f}\left(\vec{x}(t),\vec{v}(t)\right)\\ i(t)&=h\left(\vec{x}(t),\vec{v}(t)\right)\end{aligned} \tag{7.13}$$

Each x in the vector $\vec{x}(t)$ is a state variable of the system, and the number of state variables describes the order of the system. The observable output variable $i(t)$ is a function of the states of the system, and the external drive signal $v(t)$. This drive is also generally observable. The internal states $\vec{x}(t)$ are not observable. If the state equations are known *a priori*, the value of the output $i(t)$ can be determined for every time t. This

solution describes a time-parameterized path or trajectory in the multi-dimensional space of the state variables known as the "phase space." The observable output is a projection of this trajectory onto a single axis, the i-axis, and this describes the time-evolution of the output value $i(t)$: plotted as a function of time, this is a time series.

An example is the Lorenz system of equations, see (7.14) which are state equations in (x, y, z). The Lorenz equations were derived in 1963 as a very simplified model of convection rolls in the upper atmosphere:

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz \end{aligned} \tag{7.14}$$

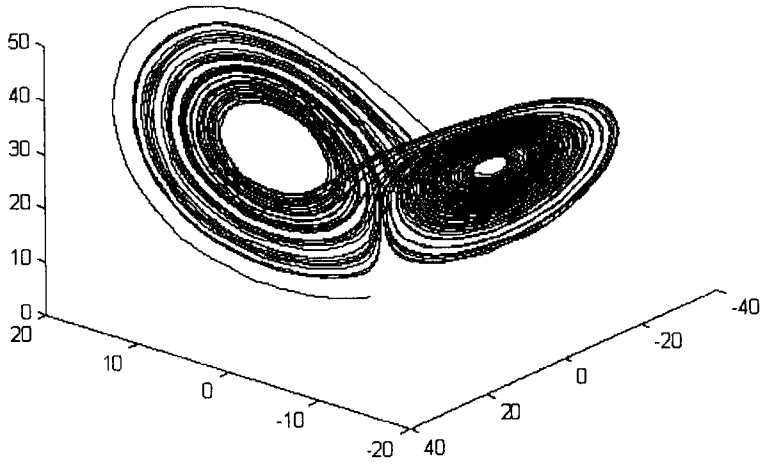


Figure 7.4 The Lorenz chaotic attractor in 3-D phase space.

The solution of this system of equations, for certain values of the parameters (σ, r, b), is a set of trajectories that never approach a stable limit cycle, yet never diverge to infinity: the trajectories describe a *chaotic attractor*, shown in Figure 7.4. Whereas the Lorenz attractor when observed in three dimensions clearly has a well-defined structure, the time-series produced by a projection onto (say) the y-axis, shown in Figure 7.5, looks completely random, and yet is completely deterministic, being defined by the Lorenz equations.

Behavioral modeling is the reverse problem to this. In our black box modeling method we start with time series data, from measurement or from simulation, and seek to develop a trajectory from this data in a suitable model phase space. We then obtain the state equations to describe this model system and implement these ODEs in the simulator. That is, we attempt to reconstruct the phase subspace (the set of reduced state variables) and a suitable flow operator (a set of differential equations on this subspace) directly from a collection of measured or simulated data. We use the data itself to infer a model.

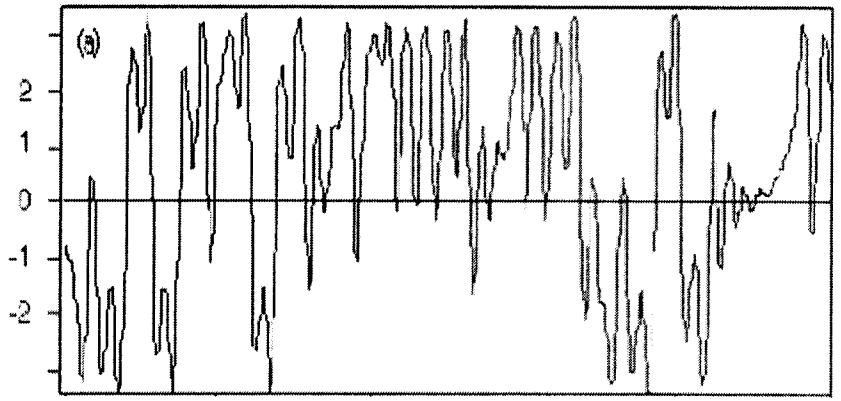


Figure 7.5 The projection of the Lorenz trajectory shown in Figure 7.4 onto the y-axis, yielding the time series.

7.3.2 Embedding

Embedding is both a simple procedure and a profound insight into the behavior of the system [10].

Given a time series of some observable $i(t)$, a trajectory in a model phase space can be constructed using a process known as embedding. A common embedding procedure is to use delayed values of the observable output. This set of quantities constitutes an embedding. The actual system trajectory (of $i(t)$) and the trajectory we create from this time-delay

embedding will differ by no more than a smooth and differential change of coordinates — the transformation relating the actual and model trajectories is a *diffeomorphism*. In other words, the trajectory in the model space preserves the dynamics of the original system. What we have done is use an observable output to retrieve explicitly the unseen internal degrees of freedom of the system — its internal state variables or their analogs. We can do this provided we take enough delays (or time derivatives, in the method described here) [7, 9].

We make use of a geometrical relationship to determine how many delays are needed. The Whitney embedding theorem states that an N-dimensional manifold can always be embedded in a Euclidean space of not more than $2N+1$ dimensions. For an example, consider a sheet of paper — a two-dimensional object. We can fold and attach the ends of the paper and get a Mobius strip, which requires three dimensions to describe it. If we then join the other ends we get a Klein bottle, which requires four dimensions to describe it. But locally, on the piece of paper, the geometry is still 2-dimensional — in the plane of the paper.

The "N-dimensional manifold" that we consider in behavioral modeling is the actual or observable system state space. The embedding theorem provides an upper bound on the number of dimensions that we will need in our model state space. The utility of this approach is that in practical cases there is an orders-of-magnitude reduction in the number of variables (dimensions) required to describe the observed behavior of the system.

How many variables will we need in the embedding? The algorithm that we use for choosing which of the dynamical variables are used for the embedding is based on the technique of false nearest neighbors [11], which can be computed using algorithms from computational geometry. We use a method described by Rhodes and Morari [12] for input-output systems. The algorithm uses the data itself to determine the optimal set of embedding variables, resulting in a compact and efficient model of vastly lower complexity than the original nonlinear system. The principle of this algorithm is illustrated in the following.

Consider the simple system comprising an observable output $i(t)$ and a single drive signal $v(t)$, which yields the response shown in Figure 7.6. Clearly the output is not a single-valued or unique function of the drive signal. For instance, the two points A and B on this curve share the same input value, but yield different outputs. So, if we sample the output $i(t)$ in terms of the drive signal $v(t)$, then both samples A and B fall into the same "bin" from v to $v+\Delta v$: the points A and B are known as false nearest

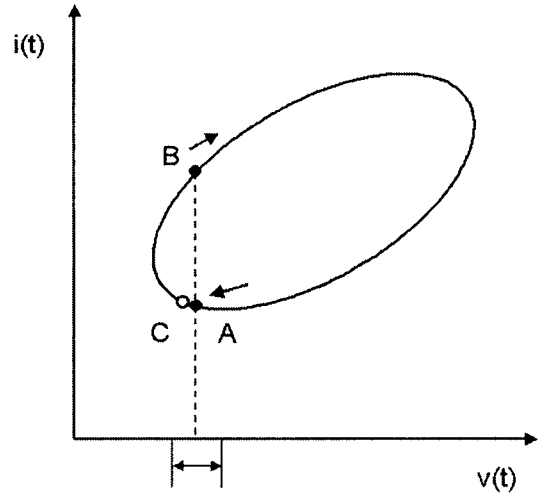


Figure 7.6 An example $i(t)$–$v(t)$ relationship: the current is not unique-valued function of the drive voltage. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

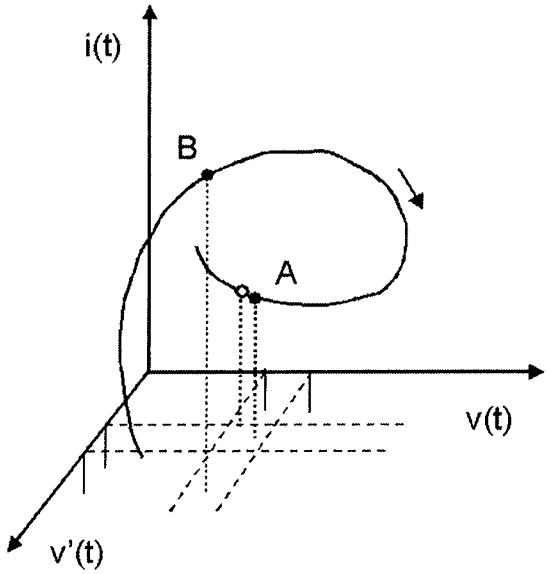


Figure 7.7 The same example function now unfolded in a higher dimensional space, of the embedding { $v(t)$, $v'(t)$ }; the current is now a single-valued functional of the embedding. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

neighbors because they are close in the input space, but are from temporally disparate locations on the response curve. The point C, which is close to point A on the response curve — a true nearest neighbor — also falls into the same bin. The number of samples in each bin will therefore be large.

We can expand the number of embedding dimensions in this example by noticing that the direction of travel around the response curve means that the first time derivative of $v(t)$ will be different at points A and B. We now plot the response $i(t)$ as a function of the drive $v(t)$ and the time derivative of the drive, and in this simple example the response curve has unfolded into a single-valued path or trajectory, Figure 7.7. Sampling $i(t)$ in the new embedding space $\{v(t), \dot{v}(t)\}$, we see that the points A and B fall into separate bins, and the true nearest neighbor point C and point A still fall into the same bin. The number of counts in each bin has fallen.

This is the basic principle of the false nearest neighbors approach. We sample the observable output variable in the embedding space, initially assuming that most of the counts will be false nearest neighbors. As we add embedding variables the state space is unfolded into higher and higher numbers of dimensions. At some point the output response curve will unfold into a single-valued trajectory, and the only points in each bin will be true nearest neighbors. If the data is sampled appropriately, this will be a small number. We monitor the density of false nearest neighbors as a function of the number of dimensions, and when this falls to a small value, this is the embedding dimension. This approach leads to fewer *ad hoc* assumptions, such as model order, compared with other recently published time-domain techniques [13, 14].

The immediate differences between this approach and our application are that we are considering a driven system, which operates over a wide bandwidth. Clearly, a constant time-delay embedding is inadequate to cover the wide timescales (bandwidth) of the excitation signal used here for the amplifier. Instead of delays, we use time derivatives of the inputs and outputs for embedding the data, yielding an expression of the general form:

$$i(t) = f\left(v(t), \dot{v}(t), \ldots, v^{(m)}(t), \ldots, i(t), \dot{i}(t), \ldots, i^{(n)}(t)\right) \qquad (7.15)$$

Equation (7.15) is a feedback model of the type represented by (7.12). These behavioral models can therefore handle systems with memory.

Equation (7.15) essentially defines an implicit, nonlinear differential equation for the behavioral model. The function f is defined by suitable multivariate approximation methods.

7.3.3 Nonlinear Function Fitting

We now have a single-valued function relating the observable output and the embedding variables. The relationship between these variables is a nonlinear one, so we seek a multivariate nonlinear function fitting method. Multivariate polynomials [15], radial basis functions (RBFs) [16], and artificial neural networks (ANNs) have been used [13, 14, 17]. ANNs are preferred, due to their asymptotic properties (i.e., they don't diverge like polynomials do) and because they give very smooth results for approximating discrete measured and simulated data. ANNs and their application to nonlinear behavioral modeling are described in detail in Chapter 6. Here we outline the procedures that we followed in our modeling example.

We use the basic structure shown in Figure 7.8. The inputs are connected to the nonlinear processing units through a set of linear weights. The nonlinear units sum all their inputs, and produce an output when this sum is above a certain threshold, which can be adjusted by the bias. The transfer function for the processing units is a "sigmoid" function — hyperbolic tangent. The nonlinear behavior is captured in these functions. The outputs from all the processing units are summed through weights at the output.

A fundamental mathematical attraction for using ANNs is found in the universal approximation theorem [18] which states that, given enough neurons in the hidden layer, a neural network of the form shown can approximate any continuous, bounded function to any accuracy that we care to specify.

Another feature of ANNs is generalization — the ability of a suitably trained network to predict correctly a response to a (set of) inputs that it has not seen before. In some cases the network can be trained to fit the target data extremely well, but performs poorly on other data of a similar class — the network has "memorized" the target data and generalizes poorly. This is a symptom of overtraining.

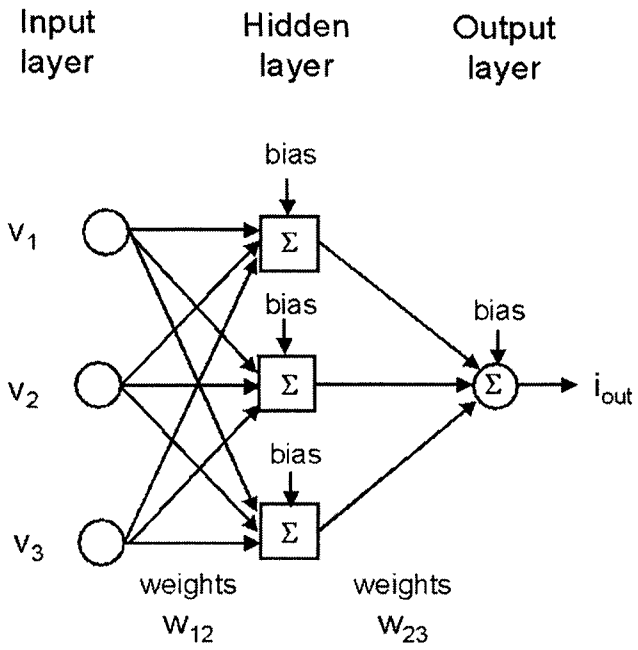


Figure 7.8 Artificial neural network structure with a single hidden layer.

Key to the design of the ANN for function approximation is the number of neurons in the hidden layer. Since the numbers of inputs and outputs are fixed — the former by the embedding procedure — the number of hidden layer neurons determines the number of weights that must be optimized during the training process to obtain the best function approximation. The values of the weights are obtained through backpropagation, a procedure where the network neuron outputs are used to update the neuron input weight values through a minimization algorithm [19]: the Levenberg-Marquardt nonlinear optimization algorithm is used. This minimization proceeds from the output of the network — comparing this with the target training value — to the input, hence the terminology. The mean square error — the difference between the ANN output and the target value — is often used as a measure of the quality of the function approximation.

Techniques for improving the generalization of the ANN include regularization [19], and early stopping [19, 20]. Regularization uses other measures in the error term that is to be minimized, such as the sum of the squares of the weight values, to reduce the potential for over-training of the network. Early stopping techniques use an additional set of data, a validation set, against which the ANN output is compared. If the error on the validation set begins to rise, even though the error on the training data set continues to fall, then this is an indication of a loss of generalization, and is an indicator to stop training the network. Cross-validation techniques [19, 20] use multiple data sets for training and validation, to improve generalization using early stopping criteria. Bayesian techniques have been employed for weight selection to improve regularization [21], and these can be used to indicate the optimum network size. Early stopping and Bayesian regularization methods work well on large ANNs, though this does not lead to parsimonious models, and may lead to poor convergence in a simulator environment.

7.4 NLTSA BEHAVIORAL MODELING PROCEDURE

The modeling procedure is outlined in flow diagram of Figure 7.9. This illustrates the general flow of activities that we need to perform to extract a behavioral model, from either measured or simulated data [13, 17], and falls broadly into three activities:

- The generation of data by simulation or measurement, which requires the design of suitable excitation signals;
- The system identification and function fitting to generate the behavioral model;
- The implementation and validation of the model in the target simulator.

7.4.1 Excitation Design

While the modeling methodology so far described is seductive in its promise to retrieve explicitly the internal dynamics of a system from an observable output, it is necessary to apply a rich enough set of signals to the system so that these dynamics are excited, and are there to be retrieved. The goal here is to choose (a set of) input signals such that all the

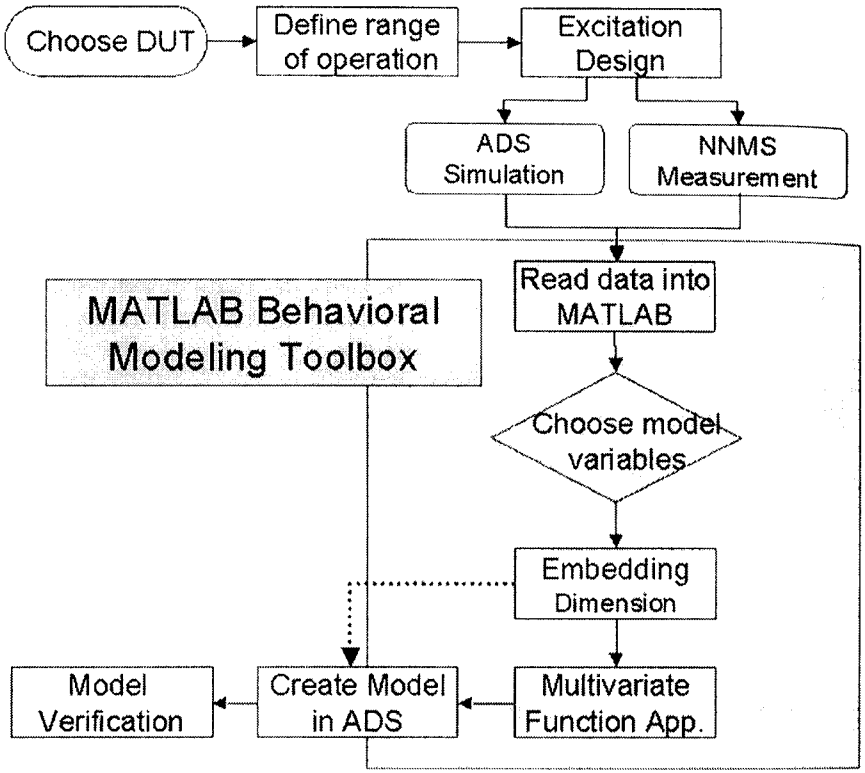


Figure 7.9 NLTSA black-box modeling procedure. The shaded region identifies those functions that have been created using MATLAB. A suitable interface between MATLAB and the Agilent-EEsof ADS simulator has been devised to generate the SDD instance, enabling the nonlinear model to be described in the simulator. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

observable nonlinear behavior of the device in a typical application is exercised. It is not necessary to design a signal that will excite all possible internal states in the device: indeed, some of these states will not be observable under the conditions of use, and in general we do not know what the internal state space of the device is. Only the observable nonlinearities are of interest in the modeling process.

A further objective relates to generalization: the excitation must include enough conditions to cover the typical operating range, as well as state or

phase space of the device or IC. We can then expect to be able to predict the behavior in this space, even if we have not measured directly under those signal conditions. The operating range of the DUT is generally specified in a datasheet. Typically we will try and cover the power and frequency span of the DUT.

We have used several classes of excitation signals for the extraction of the behavioral models. These excitations include a single-tone power and frequency sweep, which is the simplest excitation. We have augmented this by applying an additional signal at the output — this is similar to a hot-S_{22} measurement — in an attempt to excite the large signal dynamics at the output port. This can also be thought of as a synthetic load-pull experiment, so the amplifier sees a range of load conditions, and the resulting model is more likely to be transportable. We have used this excitation successfully for modeling transistors and simple amplifiers [15, 22].

We have also added a further tone at the input. This can, in principle, excite low frequency dynamics and permit study of long-term memory effects. We have used this "two-source-two-tone" excitation as a means of covering the state space of the amplifier more efficiently than with a single-tone excitation. These signals generate multiple internal states due to the nonlinear behavior of the transistors in the IC, and hence a significant number of observable states from which the model can be created.

Other excitation signals can be used, including multitone inputs and modulated signals using simple FM or complex CDMA modulation [15].

The model data can be generated either by simulation of the transistor-level circuit of the amplifier, or by direct measurement of the time-domain waveforms using a large-signal vector network analyzer (LSNA) [15, 23]. As stated earlier, the attraction of measurement-based modeling is that a low-order model of a complex component can be derived using this methodology, without knowledge of the internal circuitry or topology of the component.

A general theory of excitation design and optimal excitation design in this framework is still an open question. To date, we often apply an interactive procedure where we first identify an embedded phase space for model order reduction and then examine the coverage (by examining the density of the excitation signals) on this phase space. We can then modify our excitations signals in an attempt to provide a relatively uniform coverage of the phase space.

7.4.2 Behavioral Modeling Toolbox: System Identification and Function Fitting

The measured or simulated stimulus and response data is imported into a prototype MATLAB-based behavioral modeling toolbox, where we perform the modeling procedures.

We typically use the DUT terminal voltages and currents and their time derivatives up to second order or more as the candidate variables from which to build the models. Typically, reported methods for building an embedded phase space from a nonlinear time-series usually assume that there is a single input and single output, that the system can be described by a single characteristic timescale, and that new variables are created by delays [7]. Identifying which variables to use in the model is not a problem; a unique set of model variables is created from the delayed embeddings. Our problem differs in that we have multiple signals, and we have chosen to use time derivatives as candidate embedding variables, to enable us to describe the wide frequency range covered by the DUT. From this candidate set of model variables we need to select a subset from which to build a deterministic model. We start by using the false nearest neighbors method outlined earlier to identify a suitable set of embedding variables from these candidates. A nearest neighbors search algorithm from the *TSToolbox* [24] is used. Time-correlated samples from the time series are excluded from the search for a given sample: only data points that are beyond a time interval that is found from the autocorrelation or mutual information of the signal are used in the search. All possible combinations of the voltages, currents, and their time derivatives are submitted to the search, and the false nearest neighbors returned as a percentage for each. For all candidate sets with a low percentage score, we fit a cubic polynomial to the nonlinear function and estimate the residual error. The most promising candidate set(s) is chosen, using compactness of the candidate set and minimum residual error as guides in this choice. This is an informal application of a minimum description length criterion [25].

Once the embedding has been identified, the nonlinear function approximation is carried out. As indicated earlier, we have tried polynomials, and radial basis function approximations, but typically use feedforward artificial neural networks. We use the MATLAB Neural Network Toolbox [26]. The embedding variables — voltages, currents and their time-derivatives — are inputs to the network. The network training is carried out using backpropagation and Levenberg-Marquardt optimization:

the training is stopped manually once the training error reaches a minimum and begins to plateau. While this often gives good results, the ANN may be less than optimal [27], finding a local rather than global minimum of the function approximation. More sophisticated training methods including regularization [19] and cross-validation [20] are under investigation.

7.4.3 Implementation of the Model in ADS

The mathematical model in MATLAB is then exported as a text file to the nonlinear circuit simulator, Agilent-EEsof Advanced Design System (ADS). A proprietary piece of software code is used to convert this file into an instance of the symbolically-defined device (SDD), as illustrated in Figure 7.10. The SDD also performs the scaling and calculates the time derivatives of the variables at each time step in the simulation. Validation of the model against measurement or simulation of the transistor-level circuit is then carried out in ADS. Accuracy and speed of simulation are figures of merit for the behavioral model: the goal of a much shorter simulation time indicating that a reduced-order model compared with the full transistor-level circuit has been created.

7.5 BEHAVIORAL MODEL OF A MICROWAVE AMPLIFIER IC

The device under test (DUT) is a wideband microwave IC amplifier: Agilent Technologies, Inc HMMC-5200. This is a dc-20 GHz, 10 dB gain amplifier with internal feedback, which is designed to be used as a cascadable gain block in a variety of microwave circuit applications. It consists of eight heterojunction bipolar transistors (HBTs) configured as a compound modified Darlington feedback pair operating in Class A. The IC also contains biasing and feedback resistors, on-chip bypass capacitors, and so forth. The IC has many nodes and connections, and hence has many internal degrees of freedom. One of the key assumptions of this methodology is that only few of these internal states are important features of the observable signal. The resulting model should therefore be quite compact.

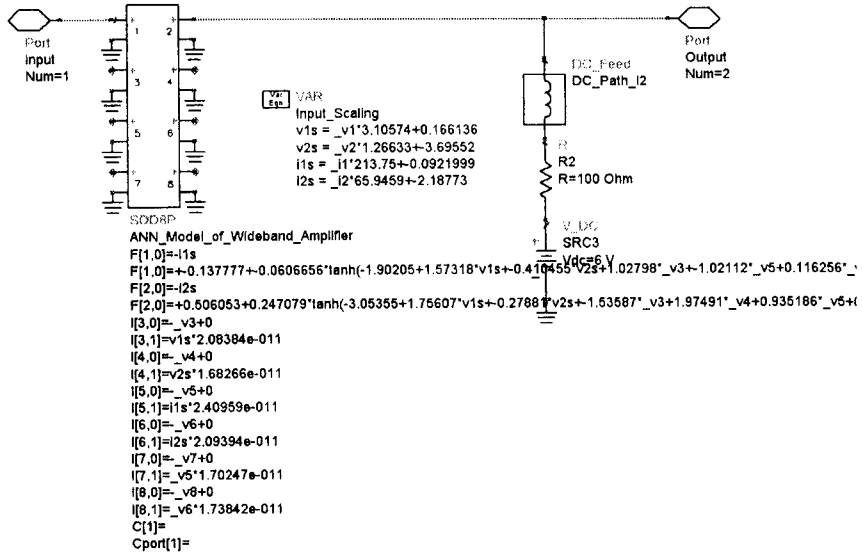


Figure 7.10 The implementation of the NLTSA behavioral model in Agilent-EEsof ADS using a symbolically-defined device (SDD). The time-derivative embedding variables are calculated "on the fly" by the SDD. The model equations for the currents are expressed as neural network expressions, at ports 1 and 2.

The power bandwidth specification of the microwave amplifier was used to determine the range of powers and frequencies for the set of excitation signals that were used to produce the data for the behavioral model generation. The excitation signal applied to this DUT was two offset tones at the amplifier input, and a further tone at the output port, swept over the frequency range from 1.2 to 10.2 GHz, and from small-signal to the P-1dB compression point: 0 dBm for each tone. The signal applied at the output port was identical in power to the input tone, and therefore about 10 dB below the output generated by the IC. The tone separation used was 600 MHz, as employed in the LSNA instrument. We performed a harmonic balance simulation using ADS; nine harmonics for each individual tone, and mixing (intermodulation) tones up to nine orders were considered. The amplifier requires a dc bias to be applied through the RF output port, via a load resistor and choke (dc feed component), and also requires dc blocks

on the input and output ports. The voltages and currents, including dc contributions, were monitored at the amplifier's RF input and output ports. The harmonic balance voltage and current data were converted into time series signals using the fast Fourier transform function in ADS.

A model was created using the current (output signals) and voltage (drive signals) at both ports, and all time derivatives of up to second-order as embeddings for the port 1 and port 2 currents. While these gave small residuals in the polynomial fit, they are not necessarily the most compact embeddings: more work is being performed on the system identification aspect. Artificial neural network models, with a single hidden layer of 40 neurons, were used for the function fitting of the input and output currents for these embeddings.

The verification procedure in ADS included:

- Single-tone power sweep harmonic balance simulation, over 1–11 GHz frequency range and to at least P–1 dB; in addition to the power magnitude and phase responses, observe harmonic distortion performance;
- Two-tone power sweep simulation, over the same power and frequency range as for the single-tone test, with a tone spacing of 100 MHz;
- Small-signal (*S*-parameter) frequency response;
- Transient simulation.

In addition to the above microwave performance of the behavioral model itself, we verified its performance as a cascade of amplifiers, thereby demonstrating the suitability of this approach for creating behavioral models for use in large system simulations.

Different power levels and frequencies to those used in the data/model generation were used for validation. In addition, we investigated the limiting cases of linear or small-signal behavior, using *S*-parameter simulation, and dc behavior. Again, it is important to note that neither small-signal nor dc data were used in the model generation procedure: only large-signal RF data were used.

In Figure 7.11, the single-tone gain compression characteristic reproduced almost exactly by the behavioral model. The frequency range is 1 to 11 GHz, the operating bandwidth of the amplifier. In Figure 7.12 phase is reproduced faithfully also.

In Figure 7.13 we show the response up to the seventh harmonic for a single-tone input at 3 dBm, which is the P–1 dB compression point. There

are some deviations but this quality of prediction is not obtained from simple models or heavily-truncated (i.e., practical) Volterra-based models. The even harmonics are reproduced well here; this is not the case with simple power out–power in models which *can only* predict odd-order harmonics. Second order correctness is important, especially for long-term or slow memory effects, dc offsets, and so forth. Note also that the dc level is reproduced exactly by the behavioral model, even though no dc measurements were used in the model construction.

The model and circuit *S*-parameters are also in excellent agreement over the frequency range 1 to 10 GHz, as shown in Figure 7.14, indicating that the fully nonlinear model reduces to the correct linear behavior under small-signal conditions.

The two-tone performance of the behavioral model is also very accurate. This is shown in Figure 7.15, for fundamental input signals of 2.0 and 2.1 GHz, at 0 dBm each tone, corresponding to about 1 dB of compression.

The time-domain output voltage waveforms for the two-tone input are shown in Figure 7.16. The RF signal is modeled accurately, and the envelope signal at 100 MHz is also reproduced well. Again this is an excellent performance as no low-frequency IF signals were used in the creation of the behavioral model.

In the above verification, the behavioral model SDD and the full transistor-level circuit model were simulated in ADS. The simulation times for the circuit model and the SDD model were found to be approximately comparable: the SDD model executed in about 25% less CPU time on a PC than the full circuit model. It is expected that a compiled model with a neural network evaluation function would be significantly faster than the SDD implementation. Further, a more compact set of embedding variables could be chosen, and more compact neural network structures can be achieved by using more sophisticated training and pruning algorithms. Both strategies will also improve simulation speed.

The NLTSA behavioral model also operates successfully in transient simulation (Figure 7.17), as well as harmonic balance. The input signal level here is very high — above the P–1 dB test level. The model is predicting hard limiting behavior correctly. This is often difficult for dynamical models to predict well, especially Volterra models.

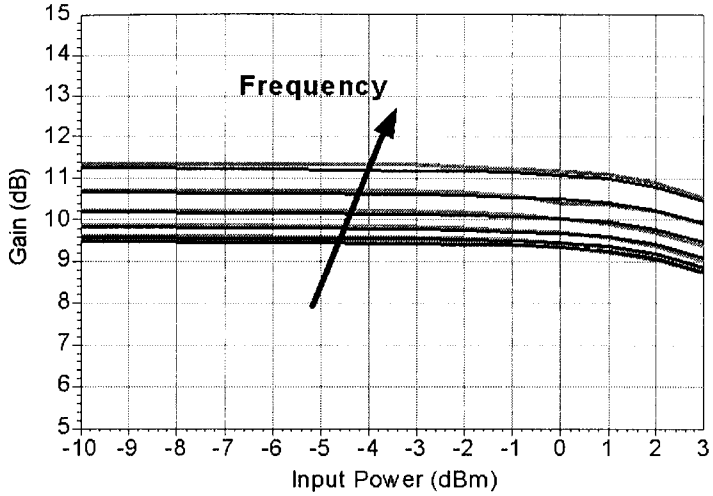


Figure 7.11 Single-tone power sweep comparing the gain compression characteristic of the NLTSA behavioral model in gray and the transistor-level circuit model in black. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

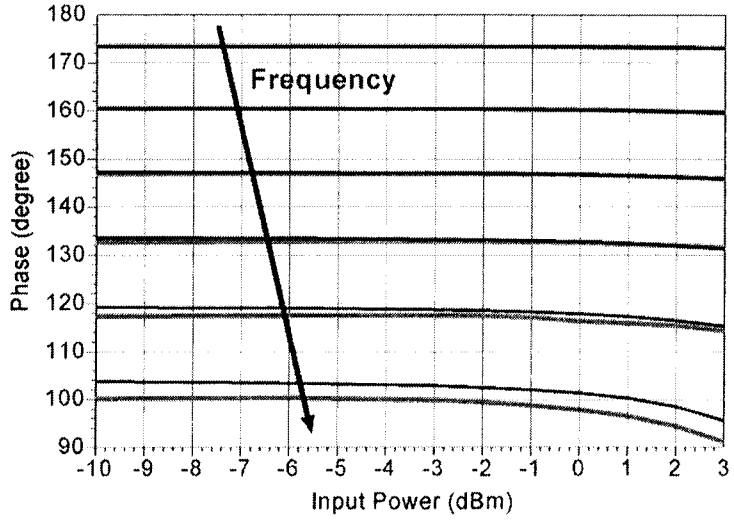


Figure 7.12 Single-tone power sweep comparing the fundamental phase response characteristic of the NLTSA behavioral model in gray and the transistor-level circuit model in black. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

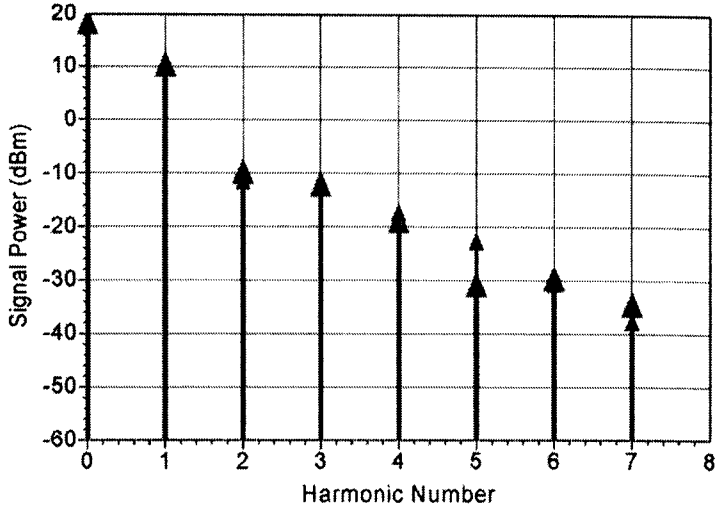


Figure 7.13 Comparison of the dc and harmonic response in HB simulation for the NLTSA behavioral model in gray, and the transistor-level circuit in black. This example is at P-1 dB compression and in the middle of the amplifier pass band, at 5 GHz. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

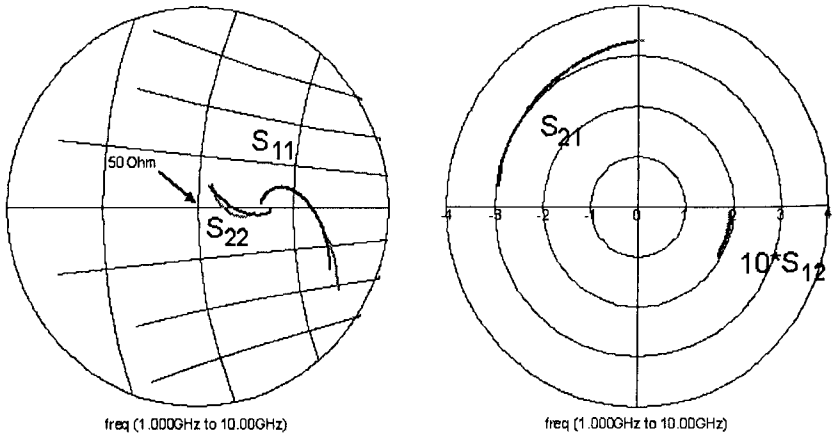


Figure 7.14 Comparison of the *S*-parameters of the NLTSA behavioral model and the transistor-level circuit, over 1 to 10 GHz. The nonlinear model reduces to linear behavior. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

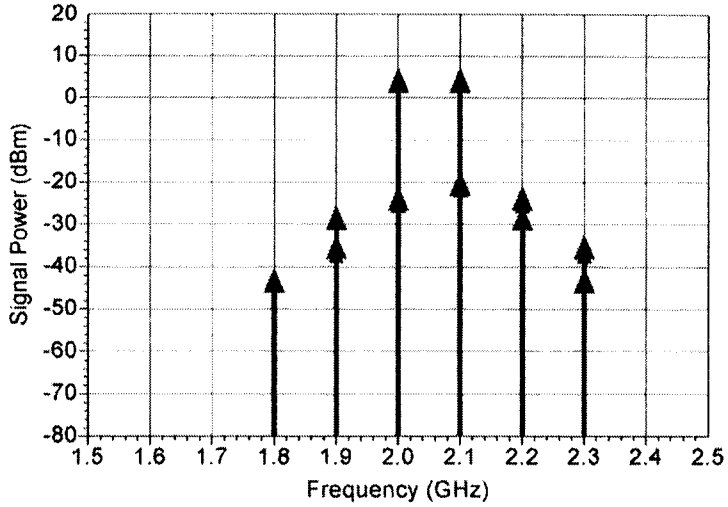


Figure 7.15 Response to a two-tone excitation at 2.0 and 2.1 GHz, comparing the NLTSA behavioral model (gray) and the transistor-level circuit (black). The input tones are each 0 dBm, giving about 1 dB compression. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

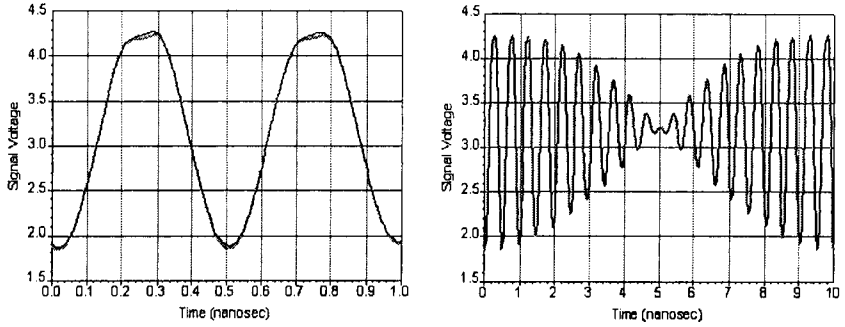


Figure. 7.16 Time-domain representation of the two-tone response of Figure 7.16, obtained by FFT. The NLTSA behavioral model (gray) displays good fidelity with the transistor-level circuit response (black), even at IF of 100 MHz. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)

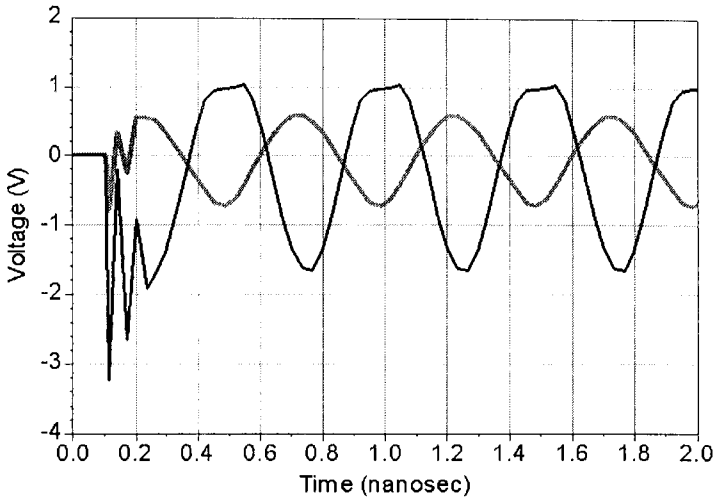
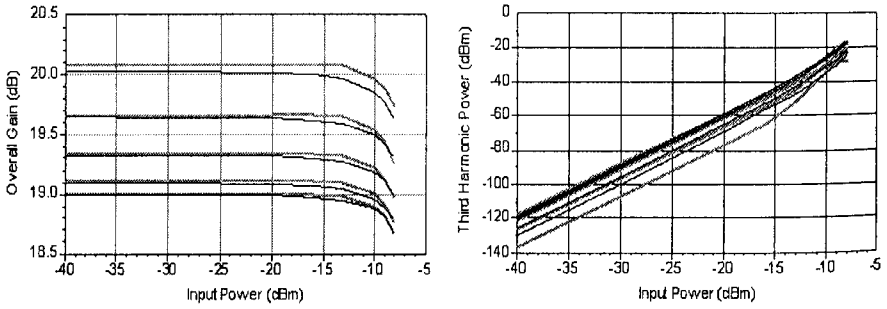


Figure 7.17 Transient response of the NLTSA behavioral model, showing rapid convergence even at a signal level of +5 dBm, well into saturation for this amplifier and much higher than the level of the training signals. (*From*: [13]. © 2004 IEEE. Reprinted with permission.)



(a) (b)

Figure 7.18 Comparison of (a) the gain compression characteristic, and (b) the third harmonic power, for a cascade of two circuits. The cascade of NLTSA behavioral models is in gray, and the transistor-level circuits in black.

While the accurate agreement between the behavioral model results and those from the transistor-level circuit are an essential first step in validating the behavioral model, the usefulness of this model is pertinent in the simulation of a module or system containing several components. To demonstrate that the behavioral model of the amplifier can be used in a system-level simulation, we place two models in cascade, and compare the results with two transistor-level circuit models in cascade. The simulation results for the gain compression characteristic are shown in Figure 7.18(a). Excellent agreement between behavioral model and transistor-level circuit is observed, and similar results are found for the harmonic performance of the cascade, as shown in Figure 7.18(b).

7.6 CONCLUSIONS

We have presented a new, general, and systematic time-domain methodology for generating nonlinear behavioral models, based on well-established techniques from nonlinear dynamics, system identification, and computational geometry. The modeling technique we have described is general, systematic, and scalable. The order of the model is contained in the embedding dimension and the ANN structure, and is vastly smaller than the number of internal degrees of freedom of the DUT.

A prototype behavioral modeling toolbox has been developed in MATLAB, that reads measured or simulated time-domain data and generates a model file that can be imported into the Agilent-EEsof ADS nonlinear microwave circuit simulator. With this toolbox we have generated a behavioral model from simulated data using a transistor-level circuit model of a wideband microwave IC amplifier. The behavioral model faithfully reproduces the circuit model electrical behavior in a wide range of validation exercises, including single-tone and two-tone power-frequency sweeps over the operating space of the amplifier, DC conditions, and *S*-parameter simulation. The cascading of two microwave amplifiers is also modeled accurately, indicating that these behavioral models can be used in system-level simulations of modules containing several amplifiers.

ACKNOWLEDGMENTS

Several colleagues have worked directly or indirectly towards the success of this work. Dominique Schreurs visited for several months and helped get us up to speed with the LSNA for practical modeling measurements. Alex Cognata made measurements of the wideband amplifier using the LSNA. Alex Pekker was an intern who helped us to establish the use of Neural Networks for the nonlinear model fitting. Jan Verspecht and Marc VandenBossche developed the LSNA instrument at the Nonlinear Measurement and Description Group of Agilent Technologies, and helped us with using and developing this instrument. Thanks are due to Gayle Collins for nonlinear dynamics and control background, and Hassan Tanbakuchi for support and instrument development. Many thanks are due to Agilent Technologies management for their continued support of this work.

References

[1] K. S. Kundert, and A. Sangiovanni-Vincentelli, "Simulation of Nonlinear Circuits in the Frequency Domain," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-5, 1986, pp. 521–535.

[2] D. Sharrit, "A New Method of Analysis of Communication Systems," *IEEE MTT-Symposium Workshop 'Computer-Aided Design of Integrated Circuits and Systems,'* San Francisco, CA, June 1996.

[3] E. Ngoya, and R. Larcheveque, "Envelope Transient Analysis: a New Method for the Transient and Steady State Analysis of Microwave Communication Circuits and Systems," *IEEE Int. Microwave Symp. Digest*, San Francisco, CA, June 1996, pp. 1365–1368.

[4] D. M. Walker, R. Brown, and N. B. Tufillaro, "Constructing Transportable Behavioral Models for Nonlinear Electronic Devices," *Phys. Lett. A*, Vol. 255, 1999, pp. 236–242.

[5] D. E. Root, J. Wood, and N. Tufillaro, "New Techniques for Nonlinear Behavioral Modeling of Microwave/RF ICs from Simulation and Nonlinear Microwave Measurements," *Design Automation Conf. Proceedings*, Anaheim, CA, June 2003, pp. 85–90.

[6] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., New York: John Wiley & Sons, 1981.

[7] H. Kantz, and T. Schreiber, *Nonlinear Time Series Analysis*, New York: Cambridge University Press, 1997.

[8] M. Casdagli, "A Dynamical Systems Approach to Modeling Input-Output Systems," in *Nonlinear Modeling and Forecasting, SFI Studies in the Science of Complexity*, Vol. XII, M. Casdagli, S. Eubank, (eds.), Reading MA: Addison-Wesley, 1992.

[9] J. Stark, "Delay Embeddings and Forced Systems," *J. Nonlinear Science*, Vol. 9, 1999, pp. 255–332.

[10] N. Gershenfeld, *The Nature of Mathematical Modeling*, New York: Cambridge University Press, 1999.

[11] M. B. Kennel, R. Brown, and H. Abarbanel, "Determining Embedding Dimension for Phase-Space Reconstruction Using a Geometrical Construction," *Phys. Rev. A*, Vol. 45, 1992, pp. 3403–3411.

[12] C. Rhodes, and M. Morari, "False-Nearest-Neighbors Algorithm and Noise-Corrupted Time-Series," *Phys. Rev. E*, Vol. 55, 1997, pp. 6162–6170.

[13] J. Wood, D. E. Root, and N. B. Tufillaro, "A Behavioral Modeling Approach to Nonlinear Model-Order Reduction for RF/Microwave ICs and Systems," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-52, No. 9, 2004, pp. 2274–2284.

[14] J. Xu, M. C. E. Yagoub, R. Ding, and Q. J. Zhang, "Neural-Based Dynamic Modeling of Nonlinear Microwave Circuits," *IEEE Int. Microwave Symp. Digest*, Seattle, WA, June 2002, pp. 1101–1104.

[15] D. Schreurs, J. Wood, N. B. Tufillaro, D. Usikov, L. Barford, and D. E. Root, "The Construction and Evaluation of Behavioral Models for Microwave Devices Based on Time-Domain Large-Signal Measurements," *IEDM Conf. Digest*, San Francisco, CA, Dec. 2000, pp. 819–822.

[16] D. M. Walker, N. B. Tufillaro, and P. Gross, "Radial Basis Model for Feedback Systems with Fading Memory," *IEEE Trans. on Circuits and Systems I*, Vol. CAS-48, 2001, pp. 1147–1151.

[17] J. Wood, and D. E. Root, "The Behavioral Modeling of Microwave/RF ICs Using Non-Linear Time Series Analysis," *IEEE Int. Microwave Symp. Digest*, Philadelphia, PA, June 2003, pp. 791–794.

[18] G. Cybenko, "Approximation by Superposition of Sigmoidal Functions," *Mathematics of Control, Signals and Systems*, Vol. 2, 1989, pp. 303–314.

[19] S. Haykin, *Neural Networks*, Upper Saddle River, NJ: Prentice-Hall, 1999.

[20] L. Prechelt, "Automatic Early-Stopping Using Cross Validation: Quantifying the Criteria," *Neural Networks*, Vol. 11, 1998, pp. 761–767.

[21] F. Dan Foresee, and M. T. Hagan, "Gauss-Newton Approximation to Bayesian Learning," *Proc. Int. Joint Conf. on Neural Networks*, 1997, pp. 1930–1935.

[22] D. Schreurs, J. Wood, N. Tufillaro, L. Barford, and D. E. Root, "Construction of Behavioral Models for Microwave Devices from Time-Domain Large-Signal Measurements to Speed Up High-Level Design Simulations," *Int. J. of RF & Microwave CAE*, Vol. 13, No. 1, 2003, pp. 54–61.

[23] J. Verspecht, P. Debie, A. Barel, and L. Martens, "Accurate On-Wafer Measurement of Phase and Amplitude of the Spectral Components of Incident and Scattered Voltage Waves at the Signal Ports of a Nonlinear Microwave Device," *IEEE Int. Microwave Symp. Digest*, June 1995, pp. 1029–1032.

[24] *TSToolbox*, http://www.physik3.gwdg.de/tstool/index.html.

[25] K. Judd, and A. Mees, "On Selecting Models for Nonlinear Time Series," *Physica D*, Vol. 82, 1995, p. 426.

[26] http://www.mathworks.com.

[27] M. Small, and C. K. Tse, "Minimum Description Length Neural Networks for Time-Series Prediction," *Phys. Rev.* E 66 066701, 2002.

About the Authors

Nuno Borges Carvalho received his diploma and doctoral degrees in electronics and telecommunications engineering from the University of Aveiro, Aveiro, Portugal, in 1995 and 2000, respectively. From 1997 to 2000 he was an assistant lecturer at the same university. Currently, he is an auxiliary professor at the university and a senior research scientist at the Telecommunications Institute. His main research interests include CAD for nonlinear circuits and design of RF-microwave power amplifiers. He was the recipient of the 1995 University of Aveiro and the Portuguese Engineering Association Prize for the best 1995 student at the University of Aveiro, the 1998 Student Paper Competition (third place) presented at the IEEE International Microwave Symposium, and the 2000 IEE Measurement Prize. He has been a reviewer for several magazines, including *IEEE Transactions on Microwave Theory and Techniques*. He is a member of the IEEE.

Edouard Ngoya received his Ph.D. degree in electronics from the University of Limoges in 1988. Since 1990 he has been a senior researcher of the French Centre National de la Recherche Scientifique (CNRS) at IRCOM University of Limoges, where he has initiated key circuit simulation and modeling techniques and contributed to the development of several CAE tools for nonlinear RF and microwave circuits. He is the author of the GoldenGate simulation tool from Xpedion Design Systems. His current domain of interests include analog system-level modeling and large-scale RF and microwave circuit simulation techniques.

José Carlos Pedro received his diploma and doctoral degrees in electronics and telecommunications engineering from the University of Aveiro, Portugal, in 1985 and 1993, respectively. From 1985 to 1993 he