New Behavioral-Level Simulation Technique for RF / Microwave Applications. Part II: Approximation of Nonlinear Transfer Functions

S. L. Loyka[1,2] J. R. Mosig[3]

[1]LACIME-ETS, Ecole de Technologie Superieure, 1100, Notre-Dame St. West, Montreal (Quebec), H3C 1K3, Canada; e-mail: segey.loyka@ieee.org
[2]Belorussian State University of Informatics and Radioelectronics, P. Brovki St. 6, Minsk 220027, Belarus
[3]Swiss Federal Institute of Technology, LEMA-EPFL, Ecublens, CH-1015 Lausanne, Switzerland; e-mail: juan.mosig@epfl.ch

Received 17 March 2000; revised 3 April 2000

ABSTRACT: The instantaneous quadrature technique is an efficient tool for nonlinear behavioral-level simulation of RF / microwave circuits or systems over wide frequency and dynamic ranges. In order to obtain accurate simulation results, accurate approximation / representation of the nonlinear transfer functions (or factors) as well as accurate measurement (or circuit-level simulation) of the amplitude (AM–AM) and phase (AM–PM) nonlinearities are required. In this paper, we consider how to approximate these transfer functions (factors) using splines, orthogonal and nonorthogonal series expansions, and evolutionary programming techniques (genetic algorithm and neural networks) with viewpoint of the simulation accuracy. The influence of AM–AM and AM–PM measurement (or simulation) inaccuracy and noise on the entire simulation accuracy is also discussed. Series expansion methods are proposed as a tool to filter out the measurement noise. © 2000 John Wiley & Sons, Inc. Int J RF and Microwave CAE 10: 238–252, 2000.

Keywords: behavioral-level simulation; nonlinear transfer function; approximation

1. INTRODUCTION

In the previous paper [1], we considered a new behavioral-level technique (the instantaneous quadrature technique) for nonlinear modeling and simulation of RF/microwave communication systems. This technique combines the advantages of the two well-known techniques, the quadrature modeling technique and the discrete technique, and allows one to simulate the system performance over wide frequency and dynamic ranges (i.e., to predict harmonics of the carrier frequency and even-order nonlinearities, to account for the frequency response (input/output matching networks) and, in principle, to model the bias decoupling network effect [2, 3]). However, in order to obtain accurate simulation results, one must pay much attention to the accuracy of the instantaneous transfer factors that play a very important role in the entire simulation process. In particular, one should pay much attention to the accuracy of AM–AM and AM–PM characteristic measurements (and especially for the second-order characteristics), their approximation/representation accuracy, filtering out the measurement noise, the accuracy of transferring envelope func-

Correspondence to: S. L. Loyka

tions into instantaneous ones, and the accuracy of instantaneous characteristic approximation/representation (in order to filter out the computational noise). This paper is organized as follows. Section 2 gives methods of the transfer function approximation/representation. In Section 3, these methods are validated using measured data for a microwave amplifier. In Section 4, we discuss the impact of approximation accuracy and measurement noise on the entire simulation accuracy.

2. REPRESENTATION/APPROXIMATION OF TRANSFER FUNCTIONS

In order to start simulation using the instantaneous quadrature technique, one needs the instantaneous transfer factors or functions. (By transfer function we mean the output voltage as a function of the input voltage, and by transfer factor we mean the voltage gain as a function of the input voltage.) Usually, fundamental (first-order) and second-order AM–AM and AM–PM characteristics are measured or simulated by a circuit-level simulator, and then these characteristics are transformed into the instantaneous transfer factors using the integral equation technique [1]. While doing so, one should keep in mind the following:

1. AM–AM and AM–PM measured (or simulated) data are available for a finite set of points. During simulation, these data are usually required for another set of points. Thus, a mathematical technique is required to transform these data from one set of points to another. Besides, these data must be stored in some form (look-up tables, for example) and that technique can also be used for this purpose. This is a representation problem.
2. While the importance of higher-order derivatives of elements' characteristics for the nonlinear circuit-level simulation is well recognized [4–6], the importance of higher-order derivatives for the nonlinear behavioral-level simulation is not so well understood [7]. Let us now consider this issue in more detail. Measured or circuit-level simulated AM–AM and AM–PM data contain not only the "real" characteristics (without any disturbance), but also measurement (or simulation) noise and distortions due to instrument (or simulator) noise and inaccuracy. These distortions and noise may severely degrade the entire simulation accuracy due to a nonlinear character of the problem. This can be illustrated by the following argument. Let us assume that the time-domain transfer function is known with some inaccuracy:

$$y(x) = F_0(x) + \Delta F(x), \tag{1}$$

where y is the output voltage, x is the input voltage, F_0 is the "real" transfer function, and ΔF is a disturbance. Using (1), we calculate the output signal in the time domain [(for given $x(t)$] and then transform it to the frequency domain. Obviously, the output spectrum contains two components, the "real" spectrum and the disturbance, due to the linearity of the Fourier transform:

$$S_y = S_0 + \Delta S, \tag{2}$$

where

$$S_0 = \mathbf{FT}(F_0), \qquad \Delta S = \mathbf{FT}(\Delta F), \tag{3}$$

and **FT** is the Fourier transform. We expect that the analysis dynamic range is at least 100 dB [8]. Thus, making a rough estimation, one obtains:

$$\Delta S < 10^{-5} S_{0,\max} \qquad \text{and} \qquad \Delta F < 10^{-5} F_0, \tag{4}$$

where $S_{0,\max}$ is the maximum level in the "real" spectrum. Hence, if we wish to obtain the output spectrum within the 100-dB dynamic range in a general case, the relative disturbance level of the transfer function must be not higher than 10^{-5}—much smaller than a typical measurement error. Fortunately, this consideration says nothing about the specific spectrum of the distortion. For example, if ΔF = constant, then the output spectrum has no distortions except at $f = 0$. One usually does not consider the DC component in system-level RF/microwave applications and, hence, this spectrum can be considered to be undis-

torted. The next example is $\partial \Delta F / \partial x =$ constant. In this case, we just predict the transfer gain with an error; all nonlinear products are predicted correctly (within the accuracy of the gain prediction—all these components are just scaled by this constant). Thus, (4) can be relaxed to, say, $\Delta F < 10^{-2} F_0$ (1% accuracy). When ΔF has nonzero higher derivatives, it may generate nonlinear noise which will disturb the nonlinear products we are looking for (for example, some specific intermodulation products). Thus, it is very important to consider not only the magnitude of the transfer function distortion (error), but also its spectrum (higher-order derivatives): we can afford some higher-level disturbance in the transfer function, but only with a specific spectrum. An efficient way to analyze and to control the spectrum of the distortions is to use a series expansion of the transfer functions (the envelope as well as instantaneous ones). We should note that ΔS has usually a noise-like form of quite a small level, while S_0 represents the required spectrum plus nonlinear distortions. Hence, when ΔS is smaller than all the spectral components in S_0 that are under consideration, ΔF has not much influence on the simulation accuracy. This is the case when all the considered spectral components are of high enough level (in the large-signal regime, for example). ΔF decreases the simulation accuracy of mainly small-level spectral components.

3. Taking into account the previous item, one should "filter out" the measurement (or simulation) noise of the envelope and instantaneous transfer functions before starting simulation (we shall consider this item in detail below). Series expansion can be used for this purpose. We just decrease the expansion order in order to suppress the higher-order noise (disturbance) components. However, we should not decrease it too much because it gives rise to the approximation error (using this series, we approximate the "real" transfer function with larger error). Obviously, there is some optimum value for the series expansion order, which depends on the transfer function itself as well as on the measurement (simulation) noise level. (However, to the best of the authors' knowledge, there is no mathematical technique to find this optimum.)

In the mathematical language, the problem of transfer function (factor) representation can be formulated as follows. We know transfer functions (AM–AM and AM–PM, or in-phase and quadrature components) for a finite set of points (from measurements or circuit-level simulation):

$$y_i = F(x_i), \tag{5}$$

where y_i is the output amplitude or phase, and x_i is the input amplitude, but the simulation usually requires this function to be defined over a continuous interval (strictly speaking—for another finite set of points, but it is more convenient to consider a continuous interval because we do not know this set in advance):

$$y = F(x), \qquad x \in [-x_{\max}, x_{\max}]. \tag{6}$$

Thus, we need a method to obtain (6) using (5). Several such methods are available [9–15], which make use of

1. splines
2. Bessel series
3. sine/cosine series
4. Chebyshev polynomial series
5. power series
6. interpolating polynomials
7. evolutionary programming techniques (genetic algorithm and neural networks)
8. wavelets

Since splines are a very traditional tool [16], we shall not discuss it in detail. However, we should note that they usually provide the best accuracy (in terms of transfer function representation) [9], but do not allow one to filter out the measurement noise. Thus, the overall simulation accuracy is not very good when the measurement noise contribution is substantial. Interpolating polynomials are also widely used for such a problem [10–11], but, unfortunately, they also do not allow one to filter out the measurement noise and their accuracy is not very good. On the contrary, a series expansion gives us a good possibility to filter out the measurement noise by changing the expansion order. The approximation (representation) accuracy can also be quite good if the ex-

pansion order is chosen appropriately. Thus, we consider these methods in more detail.

2.1. The Approximation / Representation Using Bessel Series

In this case, we represent the transfer function in the following form:

$$y(x) = \sum_{n=0}^{N} a_n J_1((2n+1)\xi x), \qquad (7a)$$

$$y(x) = \sum_{n=0}^{N} b_n J_0((2n+1)\xi x), \qquad (7b)$$

where J_0 and J_1 are the zero- and first-order Bessel function of the first kind, correspondingly, and

$$\xi = \frac{\pi}{2x_{\max}}. \qquad (8)$$

Both expansions are valid over the interval $x \in [-x_{\max}, x_{\max}]$; (7a) is used for the odd part of $y(x)$ and (7b) is used for the even part. Why is this particular form of the Bessel series expansion used? First of all, we use J_1 and J_0 because they are odd and even functions, correspondingly. Some additional motivation is as follows [17]. A typical odd transfer function has a form shown in Figure 1 (solid line; this is a typical example of a saturating nonlinearity). We choose the scaling factor inside the Bessel functions in (7) in such a way that the first term resembles the transfer function and, at the same time, does not diverge substantially from it near the ends of the approximation interval (see Fig. 1, dashed line). Higher-order expansion terms correct the divergence inside the interval and have small value at the ends. If one includes the even-order terms $(2n\xi x)$ in (7), the approximation error will decrease. However, the number of terms is two times higher in this case, which substantially increases the simulation time. Comparing these two expansions (with and without even-order terms) for the same number of terms, we find that the approximation error is approximately the same. An additional argument for using only odd-order terms in (7) is that if the first-order envelope transfer function is expanded as in (7a), we can use its expansion coefficients in order to expand the instantaneous transfer function in the sine series (see eqs. (30)–(32) in [1]); thus, we do not need to solve the integral equations to transform the envelope transfer function into the instantaneous one in this case.

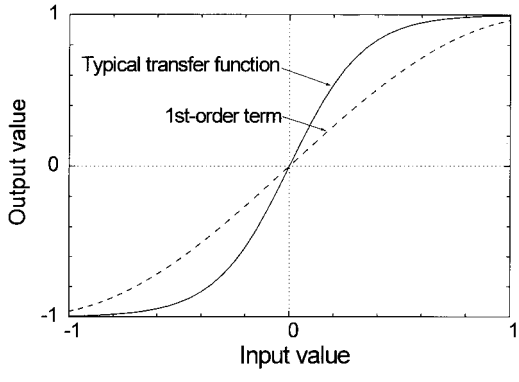

Figure 1. A typical odd input–output voltage transfer function (solid line) and the first-order expansion term in (7a) (dashed line).

We should note that the basis functions in (7) are not orthogonal ones; thus, we need some untraditional technique to calculate the expansion coefficients (a_n). The least-squares method together with the singular-value decomposition (SVD) technique can be used in the following way [16]. We calculate the expansion coefficients in such a way that the following expression is minimized (the least-squares method):

$$\varepsilon^2 = \left(\sum_{n=0}^{N} a_n f_n(x_i) - y_i \right)^2 \Rightarrow \min, \qquad (9)$$

where $\{y_i, x_i\}$ are the measured points of the transfer function [see (5)], $i = 1, M$, and M is the number of the measured points, and $f_n(x)$ are the basis functions (J_1 or J_0). Using a standard approach, we obtain from (9) the well-known matrix equation,

$$\mathbf{A} \cdot \mathbf{a} = \mathbf{b}, \qquad (10)$$

where

$$A_{in} = f_n(x_i), \qquad b_i = y_i; \qquad (11)$$

(10) is a typical system of linear equations. However, when solving this system, we may encounter the following two problems: (1) the system matrix **A** is near singular and, consequently, the accuracy of a traditional solution technique will be very poor, and (2) the number of measured points is larger than the number of expansion terms, $M > N + 1$ [the system (10) is overdetermined]. The latter problem appears because one wishes to use as many measured points as possible in order to

improve the simulation accuracy, and the number of expansion terms is limited because one needs to filter out the measurement noise (or because of the limited computer resources). Besides, it is more convenient to vary M and N independently to optimize the simulation technique for a particular problem. The former problem may appear due to a particular choice of the basis functions. In order to solve (10) under these circumstances, we can use the SVD technique [16]. We represent $\mathbf{A}$ in the form:

$$A_{in} = \sum_{k=0}^{N} w_k U_{ik} V_{nk}, \tag{12}$$

where w_k are the singular values, and the matrixes $\mathbf{U}$ and $\mathbf{V}$ are orthogonal in the sense that their columns are orthogonal:

$$\sum_{i=1}^{M} U_{ik} U_{in} = \sum_{j=0}^{N} V_{jk} V_{jn} = \delta_{kn}, \tag{13}$$

and δ_{kn} is the Kronecker's delta. A procedure to calculate $\mathbf{w}$, $\mathbf{U}$, and $\mathbf{V}$ is given, for example, in [16]. An approximate solution of (10), which minimizes the difference $|\mathbf{A} \cdot \mathbf{a} - \mathbf{b}|^2$, is given by the following expression:

$$a_n = \sum_{j=0}^{N} \sum_{i=1}^{M} \frac{U_{ij} b_i}{w_j} V_{nj}. \tag{14}$$

We can use this approach for the series expansion using not only Bessel functions, but also other nonorthogonal basis functions (power series, for example). The SVD technique performs quite well even for highly overdetermined systems (say, when $M = 100$ and $N = 10$).

In order to obtain an accurate approximation, the number of measured points M must be large enough (say, $100 \ldots 1000$). If it is not, the approximation accuracy may be poor. However, it can be improved in the following way. We build a continuous transfer function using the measured points and splines in between these points. Further, we determine another set of points $\{y_i, x_i\}$, $i = \overline{1, M^*}$, using this function, where M^* is large enough. (In this way, we practically do not lose the accuracy if a sufficient number of the measured points is used, say, if $M > 20 \ldots 50$.) Using this method, one can also locate the sample points x_i in an arbitrary way without any dependence on the measured points (of course, within the same interval). This is very important for the Chebyshev approximation (see Sec. 2.3).

2.2. The Approximation / Representation Using Sine / Cosine Series

This is a conventional approach—we just use a modified Fourier series to represent the transfer function,

$$y(x) = \sum_{n=0}^{N} a_n \sin((2n + 1)\xi x), \tag{15a}$$

$$y(x) = \frac{b_0}{2} + \sum_{n=1}^{N} b_n \cos(2n\xi x), \tag{15b}$$

where (15a) is used for an odd transfer function, and (15b) is used for an even one. The motivation for this particular form of the expansion is similar to that of the Bessel series expansion. (see Sec. 2.1) [17]. For instance, only even-order terms $[\cos(2n\xi x)]$ are used in (15b) because odd-order terms $[\cos((2n + 1)\xi x)]$ have zero value at the ends and, consequently, would not approximate correctly the transfer function (which is not equal to zero at the ends). The same argument explains why only odd-order terms are used in (15a). The expansion coefficients are calculated by the well-known formulas (the discrete form of these formulas is used):

$$a_n = \frac{2}{M} \sum_{i=1}^{M} y_i \sin((2n + 1)\xi x_i), \tag{16a}$$

$$b_n = \frac{2}{M} \sum_{i=1}^{M} y_i \cos(2n\xi x_i), \tag{16b}$$

where

$$x_i = \left(1 - \frac{M}{2}\right) \frac{2x_{\max}}{M}, \qquad i = \overline{1, M}. \tag{17}$$

If the number of measured points M is not large enough ($100 \ldots 500$) or if they are located not as in (17), one can use splines in the same way as in the preceding section. Note that the transfer functions are usually measured over the interval $x \in [0, x_{\max}]$, but, as required by (17), $x_i \in [-x_{\max}, x_{\max}]$. In order to overcome this difficulty, we assume $y(-x) = -y(x)$ for the first-order transfer function (because only the odd part of the transfer function generates spectral components within the first harmonic zone) and $y(-x)$

$= y(x)$ for the second-order one. Note also that both expansions (16) are valid over the interval $x \in [-x_{\max}, x_{\max}]$. Outside this interval, they generate a periodic continuation of the function.

2.3. The Approximation / Representation Using Chebyshev Polynomials

Chebyshev polynomials are widely used for numerical analysis due to their exceptional properties [18]. The most important ones (for our particular application) are their fast convergence (it is a common belief that Chebyshev expansions are among the most rapidly converging ones [18]) and their orthogonality over a discrete set of points (not only over a continuous interval), which gives us the possibility to use the fast Fourier transform for the expansion coefficient evaluation. The Chebyshev polynomial expansion can also be used for filtering out the measurement noise.

In the Chebyshev series expansion method, the transfer function is represented in the following form:

$$y(x) = \frac{a_0}{2} + \sum_{k=1}^{N} a_k T_K(x/x_{\max}), \quad x \in [-x_{\max}, x_{\max}], \tag{18}$$

where $T_k(x)$ are the Chebyshev polynomials (the basis functions). Note that (18) can be applied to an odd function as well as to an even one or to some combination [if one applies it to an odd function, for example, even-order coefficients will be automatically equal to zero—see (19)]. The expansion coefficients can be calculated using the polynomial orthogonality over a continuous interval or a discrete set of points [18], correspondingly:

$$a_k = \frac{2}{\pi} \int_{-1}^{1} \frac{y(x \cdot x_{\max}) T_k(x)}{\sqrt{1 - x^2}}\, dx, \tag{19a}$$

$$a_k = \frac{2}{M} \sum_{i=1}^{M} y_i T_k(x_i), \tag{19b}$$

where $x_i = \cos(2\pi i/M)$. The last equation is more convenient for computation because it includes the finite set of points; thus, we can use the measured values of the transfer functions directly (however, we must measure the transfer function nonuniformly—at x_i). It should be noted that the former equation is also useful—it can be used to estimate the coefficient aliasing effect [18] (do not confuse it with the spectrum aliasing effect—see below) and, consequently, to estimate the required number of points M. Note that sample points x_i are located in a nonuniform way. If the measured points are uniformly located or if their number is not large enough, one can use splines as in Sections 2.1 and 2.2 in order to transform these points to x_i.

When evaluating the expansion coefficients, one should keep in mind the following:

1. The number M of the measured points should be large enough that these points represent adequately the continuous function; otherwise, large errors in the simulation results are possible. For smooth functions, practical values of M equal to 10–100; for fast-changing functions, practical values equal 100–1000. A general rule of thumb is that the change in the transfer function for successive points should not be larger than 0.5–3 dB. This rule is valid for both formulas (19).
2. The number M of points is also of great importance for the expansion coefficient calculation using the discrete formula (19b) due to the coefficient aliasing effect [18]. In general, the expansion coefficients calculated by the continuous and discrete formulas are not equal:

$$\overline{a_k} = a_k - a_{2M-k} - a_{2M+k} + a_{4M-k} + a_{4M+k} - \cdots, \quad k = \overline{0, N}, \tag{20}$$

   where a_k and $\overline{a_k}$ denote the expansion coefficients calculated by (19a) and (19b), correspondingly. In order to get a good accuracy for $\overline{a_k}$, one must use large enough M, such that $a_{2M-k}, a_{2M+k}, a_{4M-k}, \ldots \ll a_k$. A particular value of M depends on a particular transfer function. A general rule of thumb is that $M \geq (3\text{–}10)N$.
3. In general, one should not use a too high-order expansion to filter out the measurement noise—the lower the order of expansion is used, the more measurement noise is filtered out. One should also take into account that too low-order expansions should not be used since the simulation accuracy would be very poor. A particular choice of the expansion order depends on the measurement noise level (measurement accu-

racy) and on fine details of the transfer function.

When one compares the Chebyshev polynomials method with the interpolating polynomials method [11, 19], it can be seen from (20) that interpolating polynomials (used for the approximation) will give higher-order coefficients with poor accuracy due to the aliasing effects because for interpolating polynomials $M = N + 1$. Due to this equation, they are also not optimal since there is no freedom in the choice of the number of points—one cannot use a large number of measured points for lower-order polynomials. The Chebyshev expansion method has this degree of freedom, which is its substantial advantage over the interpolating polynomial method.

When the number of points used for the expansion coefficient calculation is large enough, computational time can be too long. A fast Fourier transform algorithm (FFT) can be employed in this case to speed up the computation in the following way. Equation (19a) can be written in the following form [18]:

$$a_k = \frac{2}{\pi}\int_0^{\pi} y(\cos\theta \cdot x_{\max})\cos(k\theta)\, d\theta. \quad (21)$$

Sampling $y(\cos\theta \cdot x_{\max})$ for discrete points θ_i in a uniform way, we can apply FFT to calculate the expansion coefficients:

$$a_k = \mathrm{Re}[\mathbf{FFT}[\, y(\cos\theta_i \cdot x_{\max})]]. \quad (22)$$

A large number of sample points θ_i should be used in order to avoid the spectrum aliasing effect due to the FFT use. In this case, we also get a large number of coefficients (half the number of θ_i). Lower-order expansion can be obtained by omitting higher-order terms (due to the orthogonality of the Chebyshev polynomials). Thus, we need actually to compute FFT in (22) only for the required coefficients (usually, it is a small fraction of the total number).

2.4. The Approximation / Representation Using Evolutionary Programming Techniques

Two kinds of evolutionary programming techniques are currently used for the approximation of nonlinear transfer functions: the genetic algorithm (GA) and neural networks. Further, we consider in detail the use of the genetic algorithm for evaluating the series expansion coefficients [20] and its application to behavioral-level simulation.

The genetic algorithm appears to be a very effective tool for many kinds of RF problems [21–24]. GA methods differ from conventional ones in that (a) they operate on a group (or population) of trial solutions at the same time, and (b) they use stochastic operators (selection, crossover, and mutation) to explore the solution domain in search of an optimal solution. Generally, a genetic algorithm method performs the following main operations:

1. Encode the solution parameters as genes, and create a string of the genes to form a chromosome. In our case, series coefficients are used as genes.
2. Generate a starting population.
3. Assign fitness values to individuals in the population.
4. Perform reproduction through the fitness-weighted selection of individuals from the population.
5. Perform recombination and mutation to produce members of the next generation.

Steps 4 and 5 are performed until the termination criterion is met (either the number of generations or the fitness value is used as the criterion).

Although a binary-coding is used more frequently, we use here a real-coding in which real-value parameters are used as genes, because physical-problem codings work better if they resemble the parameters they model [22]. Binary tournament selection strategy is used for the reproduction due to faster operation and the absence of convergence problems [22]. In our case, the fitness function is defined to be the maximum difference between the transfer function and its approximation by the series expansion:

$$\mathrm{Fitness}(a_1 \cdots a_N) = \max\left(\left| y(x_i) - \sum_{k=1}^{N} a_k f_k(x_i) \right|\right), \quad i = \overline{1, M}, \quad (23)$$

where f_k are the basis functions. The difference is estimated at sample (measured) points x_i. In many cases, this fitness function is more appropriate for nonlinear analysis problems than, for example, the root-mean square difference (see, for

example, the considerations given in the beginning of Sec. 2). The single-points crossover is used, in which a random point in parents' chromosomes is selected and the portions of the chromosomes beyond the selected point are exchanged. The probability of crossover $p_{\text{cross}} = 0.7$ is chosen, since it is reported to be the optimal value [22]. During the mutation operation, a gene in a chromosome is changed with the probability $p_{\text{mut}} = 0.001-0.1$. Higher probabilities give faster convergence in the initial phases. Usually, p_{mut} corresponds to at most the mutation of one or two genes per chromosome and at least a few chromosomes per generation [22]. For a real-coding, the gene mutation is a small random perturbation of its initial value:

$$\hat{g} = g + 0.1_{\max} G(0,1), \tag{24}$$

where $\hat{g}$ and g are gene values after and before the mutation, correspondingly, $g_{\max}$ is the maximum value of the gene, and $G(0,1)$ is the random Gaussian-distribution number with the mean value = 0 and the standard deviation = 1. Elitist strategy is also employed (the best individual from the previous generation is inserted into the next one if the fitness value of the best individual in the next generation is lower than in the previous one). The choice of the initial population may substantially influence the GA performance. Therefore, some *a priori* knowledge should also be used.

Thus, using the genetic algorithm together with the fitness function given by (23) (or any other appropriate fitness function), we can evaluate the expansion coefficient a_k in a new way. In order to improve the computational efficiency of this method, we use the following: (a) the coefficients calculated by a conventional approximation method [see, for example, (14), (16) or (19)] are included in the initial population, giving a good starting point for the directed random search performed by the GA, (b) the initial population genes are generated using the Gaussian distribution with the mean equal to the conventional expansion coefficients (item "a"), and the standard deviation equal to 10–50% of the mean value, (c) more advanced forms of the genetic algorithm can also be used [24].

In order to demonstrate the efficiency of this method, we consider further the Chebyshev polynomials as basis functions and apply GA to evaluate the expansion coefficients. However, the use of GA is not limited to this particular kind of basis functions—any other basis functions can also be used (including nonorthogonal ones). The use of Chebyshev polynomials for the nonlinear transfer function approximation gives many advantages due to the exceptional properties of these polynomials [18], and, as with every polynomial expansion, it gives the possibility to control the spectrum expansion, avoiding in this way the spectrum aliasing effect. It is a common belief that Chebyshev polynomials give the best polynomial approximation to a given function on the minimax criterion [18]. This is exactly true when the approximation error is a polynomial. Here, we show that the GA method allows one to build a better polynomial approximation than the Chebyshev one on the minimax criterion. For this purpose, we used the instantaneous transfer function of a single-stage microwave amplifier, which has been calculated from the measured first-order envelope transfer function using the integral equation approach [1]. Further, we calculated the expansion coefficients a_k using the conventional approach [eq. (18b)] and the GA method described above. Since the performance of the genetic algorithm depends substantially on the initial population generated, we used some physical observations in order to improve the algorithm performance. In particular, we used the fact that the expansion coefficients calculated by the conventional method [eq. (19b)] are quite close to the optimal ones. Hence, we used the conventional expansion coefficients in the initial population, and the maximum coefficient values were set equal to the conventional coefficients. This approach speeds up the GA method substantially. Figure 2 shows the GA flow chart diagram used for the evaluation of the Chebyshev expansion coefficients. The population size was set equal to 100 individuals. The termination criterion used was the number of iterations equal to 1000. Some particular results are shown in Table I. The main result is that the GA method gives the maximum difference between the transfer function and its approximation, which is on average 1.2–1.5 times smaller than the conventional Chebyshev expansion method. Figure 3 shows the convergence curves of the GA approximation method. As can be seen from this figure, the GA method converges at approximately 100 . . . 200 iterations. Typical computational time on a Pentium II computer is several hours (for 1000 iterations).

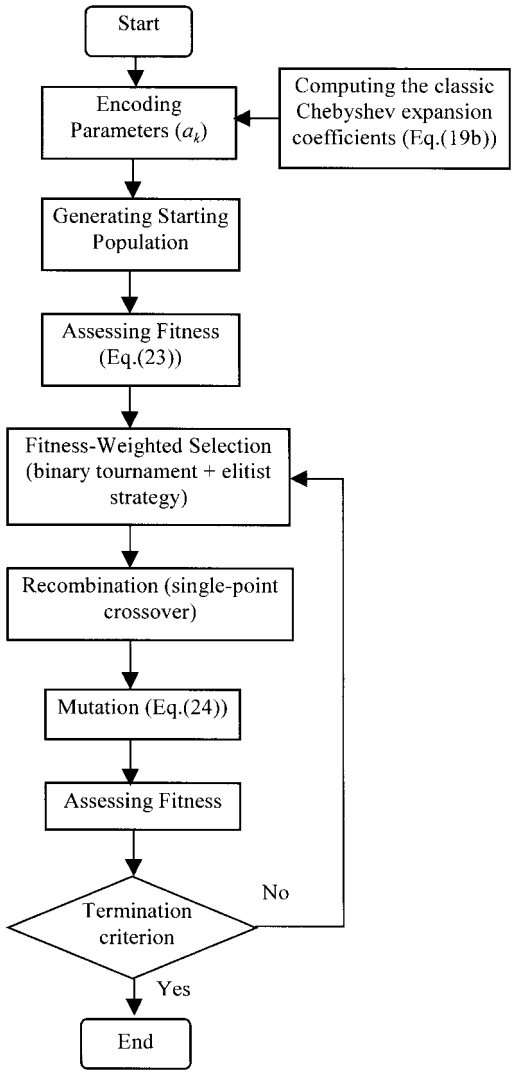

Figure 2. Flow chart diagram of the genetic algorithm employed for the evaluation of the Chebyshev expansion coefficients.

TABLE I. Comparison of the Conventional and GA Approximation Errors of the Instantaneous Transfer Function of a Microwave Amplifier

Expansion Order	Conventional Approximation Error	GA Approximation Error
4	0.034	0.029
6	0.036	0.024
10	0.015	0.0094
14	0.012	0.009

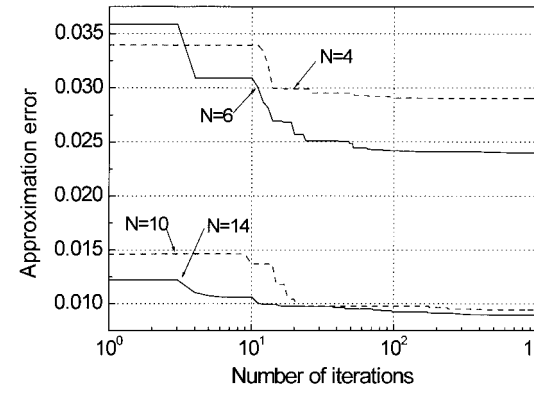

Figure 3. Typical convergence curves of the GA approximation method (approximation error versus the number of iterations).

Some additional testing of the GA method was made using a hyperbolic tangent function. (This function is frequently used in nonlinear simulation of active circuitry.) Figure 4 shows the approximation error for the conventional method and for the GA method applied to the approximation of this function. It can be seen from this figure that the approximation curve peaks for the classical Chebyshev series are not of equal amplitude. It means that this approximation is not optimal on the minimax criterion [18] and may be further improved. At the same time, all peaks of the GA method curve are almost equal, indicating that this approximation is very close to the optimal one.

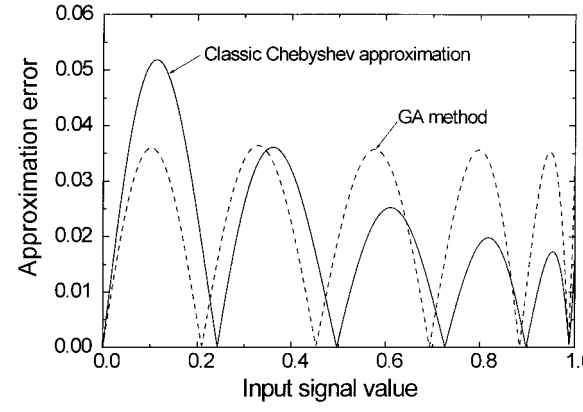

Figure 4. Approximation error of the conventional Chebyshev series (dashed line) and the GA method (solid line) for $N = 9$ (hyperbolic tangent was used as a testing function) versus input signal value.

Thus, the GA approximation method is a robust approach that gives many advantages, including optimization possibilities, global search for the optimal solution, improvement in the approximation accuracy, easy implementation of various approximation criteria, etc. We should also note that in general the GA method operates well for almost every kind of transfer function, without any strict requirements (continuous or bounded derivatives, etc.). Improvement in the approximation accuracy by 1.2–1.5 times when compared to the conventional Chebyshev polynomials approach, which is sometimes considered to be the best polynomial approximation, is possible by using the GA method. Hence, it seems that this method can potentially give the best possible polynomial approximation.

Another evolutionary programming technique, which is extensively used for RF/microwave applications, is neural networks [25–30]. The use of neural networks for function approximation is based on the universal approximation theorem, which generalizes approximations by finite Fourier series [25]. Multilayer neural networks trained by the odd and even backpropagation algorithms were used in [26] for approximating transfer functions. It was demonstrated that neural network models provide better accuracy than the classical analytical models of traveling-wave tube amplifiers. However, it has not been demonstrated that the neural network models perform better than some numerical models, for example, the series expansion models discussed above. Thus, this issue requires further investigation.

3. VALIDATION OF THE APPROXIMATION METHODS

In order to validate the approximation methods given above, the first and second order envelope transfer functions (AM–AM) of a MMIC amplifier were measured (we do not consider AM–PM in this example due to its small values) and transformed into the instantaneous transfer factors using the integral equation technique [1]. Figure 5 shows the first-order voltage transfer factors (gains) and Figure 6 shows the second-order ones. Note that the instantaneous transfer factors usually have more complex form (more oscillations, more rapid changes, etc.) than the envelope ones. Hence, it is more difficult to approximate the instantaneous factors than the envelope ones. Below, we consider the approximation of the instantaneous factors to test approximation methods using real-world data. Figure 7 shows the maximum normalized (to the factor maximum value) approximation error of the first-order instantaneous transfer factor approximated by Bessel, Cosine, and Chebyshev series. The Bessel series gives the best approximation (in terms of approximation accuracy) for $N < 20$, and the cosine series gives the best approximation for $N > 20$. The Chebyshev series gives the worst approximation in terms of maximum approximation error. However, it has the advantage that it is a polynomial approximation, and, consequently, gives a finite spectrum expansion. Hence, it allows one to control the spectrum expansion (by choosing an appropriate expansion order) and to avoid in this way the spectrum aliasing effect completely. None

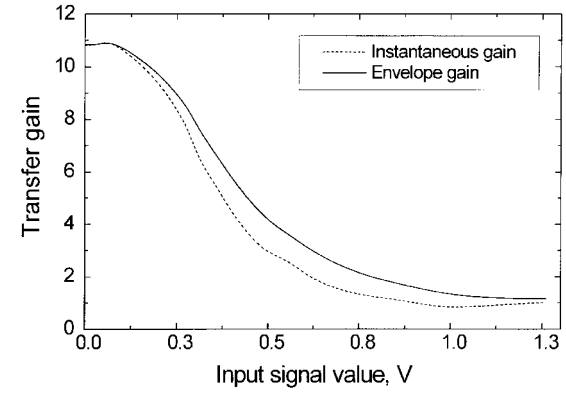

Figure 5. First-order instantaneous (dashed line) and envelope (solid line) voltage transfer gains (factors) of a microwave amplifier.

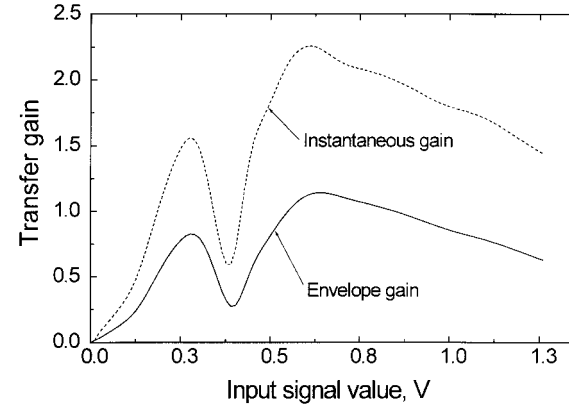

Figure 6. Second-order instantaneous (dashed line) and envelope (solid line) voltage transfer gains (factors) of a microwave amplifier.

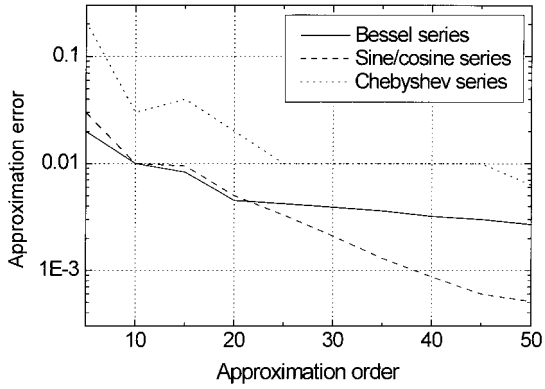

Figure 7. Comparison of first-order transfer factor approximation by Bessel, cosine, and Chebyshev series.

of the other methods has this capability: they mitigate the spectrum aliasing effect by setting high-enough sampling frequency (with same margin); however, this does not remove this effect completely and some distortions are always present over the analysis frequency range. It should be pointed out that this feature is very important when, for example, a low-noise circuit is analyzed and, consequently, high spectral purity must be maintained. The spectrum aliasing effect also limits the analysis dynamic range (due to the distortion of low-level spectral components). In general, the approximation method should be chosen keeping in mind the problem under consideration.

Figure 8 shows the approximation error of the second order transfer factor. As one may see from this figure, the Bessel and Sine series provide approximately the same approximation accuracy and the Chebyshev series provides worse accuracy. Hence, for a given approximation accuracy, the Chebyshev series will require higher approximation order (more approximation terms and, consequently, more computational time) than the Bessel and sine series. However, as it was already noted, the Chebyshev series has the advantage of providing finite spectrum expansion. One may also note that the second-order factor approximation results in higher approximation error than the first-order one or, for a given approximation error, requires a higher-order approximation.

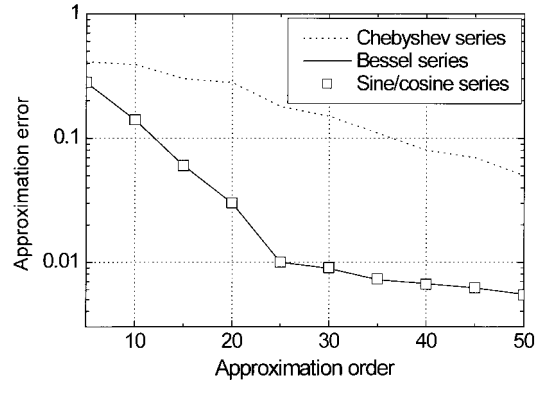

Figure 8. Comparison of second-order transfer factor approximation by Bessel, sine, and Chebyshev series.

Another important point is a dependence of the approximation accuracy on the sample point number (M). When approximating a function, one needs to know how to set M to minimize the approximation error. Figures 9–11 show the maximum normalized approximation error as a function of the sample point number. Using these

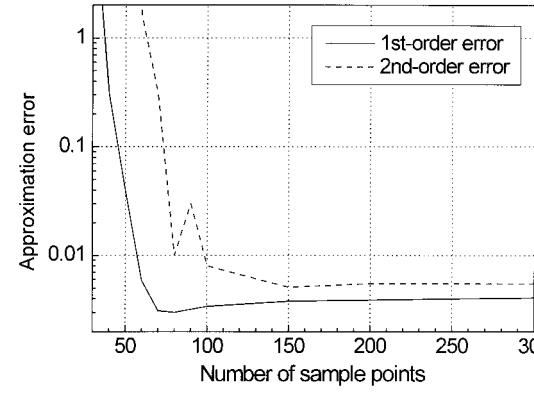

Figure 9. Normalized approximation error of the first- (solid line) and second-order (dashed line) transfer factors as a function of the sample point number (M) for the Bessel expansion; $N = 30$ and $N = 50$ for the first- and second-order, correspondingly.

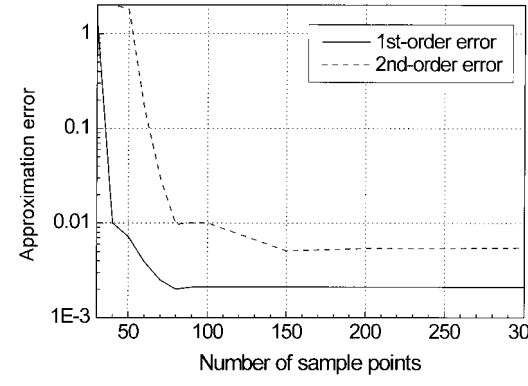

Figure 10. Normalized approximation error of the first- (solid line) and second-order (dashed line) transfer factors as a function of the sample number (M) for the sine/cosine expansion; $N = 30$ and $N = 50$ for the first- and second-order, correspondingly.

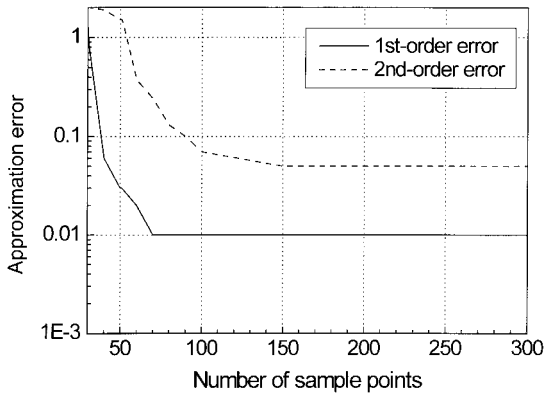

Figure 11. Normalized approximation error of the first- (solid line) and second-order (dashed line) transfer factors as a function of the sample point number (M) for the Chebyshev expansion; $N = 30$ and $N = 50$ for the first- and second-order, correspondingly.

data, we may formulate the following rule of thumb:

$$M \geq (2 \div 5)N. \tag{25}$$

It should be noted that the best approximation accuracy does not mean the best overall simulation accuracy. The overall simulation accuracy depends strongly not only on the maximum approximation error, but also on a particular approximation error "spectrum" [i.e., with what accuracy we approximate higher-order derivatives; see Sec. 2 and eqs. (1)–(3)]. This issue requires further investigation.

Inaccuracies in AM–AM and AM–PM measurement or simulation (including distortions and noise) also influence substantially the overall simulation accuracy. Careful choice of particular approximation type and order may improve the accuracy. We consider this issue in more detail in the next section.

4. APPROXIMATION ACCURACY AND MEASUREMENT NOISE

The series expansion methods presented in the preceding sections are quite efficient tools for the approximation/representation of nonlinear transfer functions for behavioral-level simulation. However, due to the nonlinear character of the problem, special attention must be paid to the approximation accuracy—even small inaccuracies in the transfer function approximation may result in very large inaccuracies in the final simulation results. The main method to control accuracy of a series expansion is through an appropriate choice of the expansion order (N): if one wishes to increase accuracy, the expansion order should be increased. Of course, there are some limitations to such increase due to the numerical stability of a computational algorithm. However, usually these limitations are quite large and do not influence substantially a practical choice of the expansion order. (For example, for the Chebyshev expansion, this limit is about 150–200, which is far beyond a practically required expansion order.)

There is a more serious limitation to the expansion order—the measurement (or simulation) noise which is always present in measured (simulated) AM–AM and AM–PM data, as shown by eq. (1). What we need is to approximate $F_0(x)$ as accurately as possible and, at the same time, to suppress $\Delta F(x)$ as much as possible.

When we consider the "spectrum" of $\Delta F(x)$ (in terms of expansion coefficients a_k), we find that it mainly contains high-order components; the contribution of lower-order components is quite small (of course, their levels depend substantially on a particular scenario). On the contrary, the "spectrum" of $F_0(x)$ contains mainly low-order components with higher-order components being much smaller. Thus, when we increase the expansion order, the accuracy of $F_0(x)$ approximation increases, but, at the same time, the noise contribution to the approximated characteristic [due to $\Delta F(x)$] also increases. This situation is illustrated by Figure 12. The optimum value of the expansion order corresponds approximately to the point where both curves intersect. Unfortunately, to the best of the authors' knowledge, there is no mathematical technique to de-

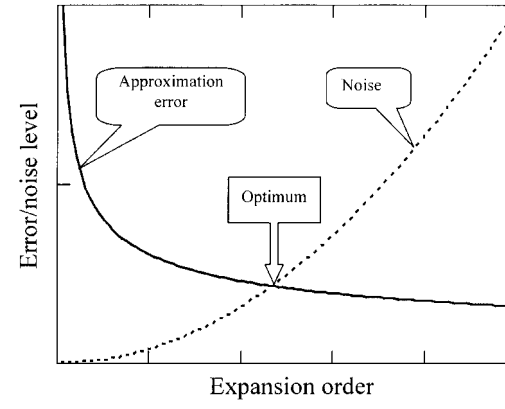

Figure 12. Approximation error (solid line) and measured noise contribution (dotted line) as functions of the expansion order.

termine the optimal order. A practical method to do this is to increase the expansion order (starting from some small value, say, $N = 5 \ldots 10$) until the simulation accuracy (for example, the accuracy of intermodulation product prediction) starts to decrease.

It should be pointed out that the simulation noise is usually much smaller than the measured one. Figure 13 gives an example of measured and HB-simulated AM–PM functions (different amplifiers were used for the measurement and simulation). As one can see from this figure, the measurement noise (ripples on the curve) is much higher than the simulation one. Thus, some measures must be taken in order to suppress the noise. One possible solution is to use during the measurements the time averaging option of the network analyzer. However, it does not allow suppressing noise completely. Besides, there is another distortion of the AM–PM data, due to the measurement instrument inaccuracy. Under these circumstances, we can use the series expansion as a tool for suppressing the measurement noise (including the measurement inaccuracy component) by controlling the expansion order. Thus, before transforming AM–AM and AM–PM measured data into the instantaneous transfer factors [1], we employ the series expansion of these data (actually, the in-phase and quadrature transfer functions are expanded). For simulated AM–AM and AM–PM data, this procedure is not required due to the small level of the simulation noise. However, special measures must be taken to ensure a high accuracy of the simulated data (for example, high relative accuracy and a large number of frequencies should be set during the HB simulation).

In order to illustrate the influence of the measurement noise on the simulation accuracy, we measured and simulated intermodulation products in the microwave amplifier considered in Section 3. Figure 14 shows the fundamental component and third to ninth order IMPs ($N = 24$; Chebyshev polynomial series was used for the instantaneous transfer function expansion). As one may see from this figure, the simulation accuracy is quite good over a wide dynamic range (in excess of 130 dB) for the optimum approximation order $N = 24$ (even seventh- and ninth-order IMPs are predicted quite accurately). However, the dependence of the simulation accuracy on the approximation order is highly nonmonotonous: increase in the approximation order over the optimum value results in an accuracy decrease rather than increase. Figure 15 gives the maximum simulation error (over the entire simulation dynamic range) for the fundamental tone and IMPs versus approximation order. The point $N = 24$ corresponds to an optimum value for almost all nonlinear products; further increase in N yields an increase in the simulation error, especially for higher-order nonlinear products. At the same time, as Figure 7 shows, the approximation error decreases monotonically (for $N > 15$) as N increases. Thus, the dependence of the simulation accuracy on the approximation accuracy is highly nonmonotonous: increase in the approximation accuracy does not guarantee an increase in the entire simulation accuracy. We attribute this fact to the measurement noise (and other meas-

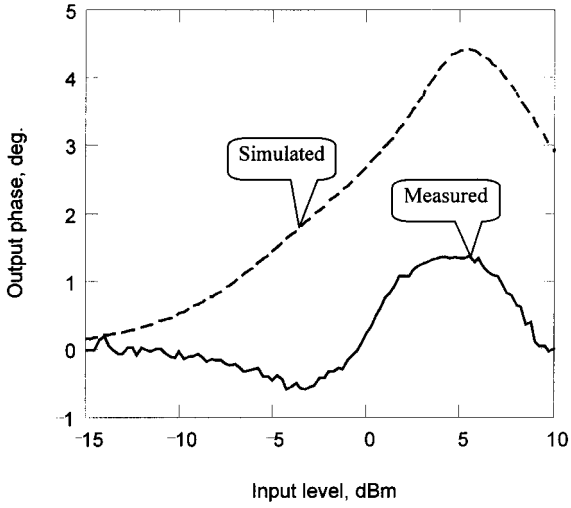

Figure 13. Measured (solid line) and HB simulated (dashed line) AM–PM characteristics (different amplifiers were used for the measurement and simulation).

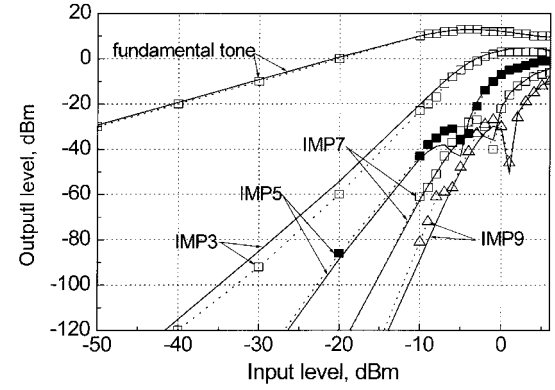

Figure 14. Measured and simulated third- to ninth-order IMPs of the microwave amplifier ($N = 24$; solid-line—simulated, squares and triangles—measured).

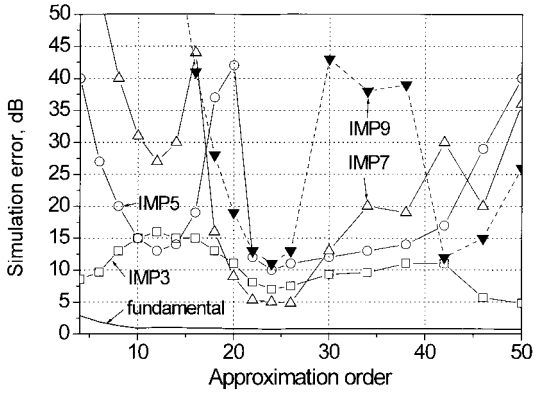

Figure 15. Maximum simulation error over the entire simulation dynamic range versus approximation order (Chebyshev series was used).

urement inaccuracies) impact as well as to the inaccuracies (computational noise, etc.) of transforming the envelope transfer functions into the instantaneous ones.

Some other factors contributing to the simulation inaccuracy are the following:

1. The bias decoupling network effect [3, 9] that is not taken into account in a conventional simulation. In principle, this effect can be modeled by the instantaneous quadrature technique.
2. The difference between static and dynamic AM–AM and AM–PM characteristics [3, 31].
3. The thermal tone spacing effect [32].
4. Memory feedback effects [3].

5. CONCLUSION

In this paper, we considered the methods of nonlinear transfer function approximation for behavioral-level simulation. Series expansion methods were proposed as an efficient tool for filtering out the measurement noise (and other measurement inaccuracies). The use of the genetic algorithm for approximation problems enables one to improve the minimax approximation accuracy over traditional techniques. The impact of the approximation accuracy and measurement noise on the entire simulation accuracy was also discussed. The approximation methods proposed were validated using measured data for a microwave amplifier. Possible topics of further investigations are the development of a general mathematical technique for finding the optimal expansion order, improvement in the GA method computational efficiency (using the gradient descent method, for example), and a detailed investigation of the impact of the transfer function disturbance (ΔF) spectrum on the entire simulation accuracy.

ACKNOWLEDGMENTS

The authors would like to thank Professor F. Gardiol, Professor V. Mordachev, Mr. J. Staudinger, and Mr. I. Cheremisinov for many helpful discussions, Mr. J. F. Zurcher for the help with measurements, and Mr. O. Kashanskii for technical assistance. The financial support of this work by the Swiss Government is also appreciated.

REFERENCES

1. S.L. Loyka and J.R. Mosig, New behavioral-level simulation technique for RF/microwave applications. Part I: Basic concepts, Int J RF Microwave CAE, this issue.
2. J. Staudinger, Behavioral analysis method applied to the design and simulation of linear power amplifiers, RAWCON'98 Workshop Modeling and Simulation of Devices and Circuits for Wireless Communications, Colorado Springs, CO, 1998.
3. W. Bosch and G. Gatti, Measurement and simulation of memory effects in predistortion linearizers, IEEE Trans Microwave Theory Techniques 37(12) (1989), 1885–1890.
4. S.A. Maas, Modeling MESFETs for intermodulation analysis of mixers and amplifiers, IEEE Microwave Theory Tech Symp Digest, 1990, pp. 1291–1294.
5. S.A. Maas, How to model intermodulation distortion, IEEE Microwave Theory Tech Symp Digest, 1991, pp. 149–151.
6. N.B. de Carvalho and J.C. Pedro, Multitone frequency-domain simulation of nonlinear circuits in large- and small-signal regimes, IEEE Trans Microwave Theory Techniques 46(12) (1998), 2016–2024.
7. S.L. Loyka, On the use of Cann's model for nonlinear behavioral-level simulation, IEEE Trans. on Vehicular Technology, 2000, accepted for publication.
8. J.F. Sevic, M.B. Steer, and A.M. Pavio, Nonlinear analysis methods for the simulation of digital wireless communication systems, Int J Microwave Millimeter-Wave CAE 6(3) (1996), 197–216.
9. J. Staudinger, Applying the quadrature modeling technique to wireless power amplifiers, Microwave J 40(11) (1997), 66–86.
10. S.L. Loyka and. V.I Mordachev, Computer-aided nonlinear simulation at the system level, Proc 5th Int Conf on Electromagnetic Interference and

Compatibility (INCEMIC'97), Hyderabad, India, 1997, pp. 93–98.
11. I.D. Cheremisinov, S.L. Loyka, and V.I. Mordachev, Synthesis of the polynomial model of nonlinear elements based on intermodulation dynamic ranges, Proc 3rd Int Conf on Telecommunications in Modern Satellite, Cable, and Broadcasting Services (TELSIKS'97), Nis, Yugoslavia, 1997, pp. 519–522.
12. A.M. Zaezdnyi, Basics of nonlinear and parametric RF Circuit analysis, Sviaz, Moscow, 1973 (in Russian).
13. P.N. Mathanov, Basics of circuit analysis: Nonlinear circuits, Vysshiaia Shcola, Moscow, 1986 (in Russian).
14. V.S. Andreev, Theory of nonlinear electrical circuits, Radio i Sviaz, Moscow, 1982 (in Russian).
15. T.K. Sarkar and C. Su, A tutorial on wavelets from an electrical engineering perspective, Part 2: The continuous case, IEEE Antennas Propagation Mag 40(6) (1998), 36–49.
16. W.H. Press, B.P. Flanneay, S.A. Teukolsky, and W.T. Vetterling, Numerical recipes in C, Cambridge University Press, 1988.
17. A.R. Kaye, D.A. George, and M.J. Eric, Analysis and compensation of bandpass nonlinearities for communications, IEEE Trans Commun 20(11) (1972), 965–972.
18. R.W. Hamming, Numerical methods for scientists and engineers, Dover Publications, New York, 1973.
19. S.L. Loyka and I.D. Cheremisinov, Validation of the high-order polynomial models used in behavioral-level simulation, 4th Int Conf on Telecommunications in Modern Satellite, Cable, and Broadcasting Services, Nis, Yugoslavia, 1999, pp. 592–595.
20. S.L. Loyka, New method for nonlinear transfer function approximation, Electron Lett 35(10) (1999), 789–791.
21. M. Johnson and Y. Rahmat-Samii, Genetic algorithms in engineering electromagnetics, IEEE Antennas Propagation Mag 39(4) (1997), 7–21.
22. D.S. Weile and E. Michielssen, Genetic algorithm optimization applied to electromagnetics: A review, IEEE Trans Antennas Propagation 45(3) (1997), 343–353.
23. S. Iezekiel and A. Feresidis, Application of genetic algorithm to sample balance analysis of nonlinear circuits, Electron Lett 34(22) (1998), 2080–2082.
24. M. Tomassini, A survey of genetic algorithms, Annual Reviews of Computational Physics, Vol. III, to be published.
25. S. Haykin, Neural networks, Prentice Hall International, New Jersey, 1999.
26. M. Ibnkahla, J. Sombria, F. Castanie, N.J. Bershad, Neural networks for modeling nonlinear memoryless communication channels, IEEE Trans Commun 45(7) (1997), 768–771.
27. Y. Harkouss, J. Rousset, H. Chéhadé, E. Ngoya, D. Barataud, J.P. Teyssier, The use of artificial neural networks in nonlinear microwave devices and circuits modeling: An application to telecommunications systems design (invited article), Int J RF Microwave CAE 9 (1999), 198–215.
28. J.A. Garcia, A. Tazón Puente, A. Mediavilla Sánchez, I. Santamaría, M. Lázaro, C.J. Pantaleón, J.C. Pedro, Modeling MESFETs and HEMTs intermodulation distortion behavior using a generalized radial basis function network, Int J RF Microwave CAE 9 (1999), 261–276.
29. M. Vai and S. Prasad, Neural networks in microwave circuit design—Beyond black-box models (invited article), Int J RF Microwave CAE 9 (1999), 187–197.
30. P. Burrascano, S. Fiori, and M. Mongiardo, A review of artificial neural networks applications in microwave computer-aided design (invited article), Int J RF Microwave CAE 9 (1999), 158–174.
31. A.A. Moulthrop, C.J. Clark, C.P. Silva, M.S. Muha, A dynamic AM/AM and AM/PM measurement technique, IEEE Int Symp on Microwave Theory and Techniques, Denver, CO, 1997.
32. P.M. McIntosh and C.M. Snowden, The effect of a variation in tone spacing on the intermodulation performance of class A and class AB HBT power amplifiers, IEEE Int Symp on Microwave Theory and Techniques, Denver, CO, 1997.

BIOGRAPHIES

Sergey L. Loyka for photo and bio see pg. 237.

J. R. Mosig for bio see pg. 237.