

MODELING MEMORY EFFECTS IN NONLINEAR SUBSYSTEMS BY DYNAMIC VOLTERRA SERIES

E. Ngoya, A. Soury

IRCOM – UMR CNRS 6615 – University of Limoges, 123 av. A. Thomas, 87060 Limoges, France
ngoya@ircom.unilim.fr

Abstract - Design challenges resulting from single chip integration of RF and microwave transceivers is requiring both more powerful analog simulation techniques and more accurate behavioral models that allow efficient hierarchical simulation. Today system level models are limited by their inability to effectively cope with nonlinear memory effects. The paper describes a modeling approach based on modified Volterra series which accounts efficiently for memory effects in nonlinear subsystems.

I. INTRODUCTION

Reduction of time to market of communications systems has put a higher demand on simulation, both at circuit and system level. It is a fact that the simulation efficiency depends primarily on the component model accuracy. However if component model accuracy can be considered satisfying today in circuit-level simulation, this is far to be the case in system-level, as regards subsystems working on nonlinear regime, especially for integrated power amplifiers and converters. The commonly used amplitude-amplitude and amplitude-phase modulation (AM-AM, AM-PM) model suffers severely from the envelope memoryless assumption, which limits its application only to very narrow band systems. Several enhancements of the AM-AM, AM-PM model have been proposed by the past [1-4] to handle memory effects. However being primarily designed for traveling wave amplifiers (TWA) their efficiency is very limited when applied to integrated circuits [5] as model prediction is poor as regards digital modulation stimuli. Recently there is a great interest symbolic model order reduction (MOR) from Kirchoff law defining the circuit at the transistor level [6-9]. These are very promising approaches, though the work reported so far succeed only for mildly nonlinear regimes. The MOR approach however suffers a basic limitation that it does not allow the model to be derived from physical measurements of the subsystem, hence preventing an escape from transistor modeling inaccuracies. This may also exhibit poor simulation speed and convergence problems as it requires the solution of a fairly large set of nonlinear equations. The model described below tries to reconcile the various aspects of simulation speed, modeling accuracy and modeling situations. It is a functional black box model approach that can be derived either from commonly

available circuit simulation tools or from common physical measurements equipments.

The model is based on a first order truncation of a modified Volterra series. Volterra series[10] has a bad reputation among engineers of being a cumbersome technique; but yet it is an effective formalism for representing nonlinear systems with memory, especially when the system is large and distributed. Nevertheless the classical form of Volterra series has poor convergence properties and in practice it is hardly possible to measure its kernels of order more than two to three. These limitations make the classical Volterra series inefficient for most nonlinear IC applications. So this paper proposes to use a modified form of Volterra series [11] which lessen the above limitations, providing fairly good accuracy with only the first order kernel. The extraction procedure of the model is then easily affordable, based on single or two-tone simulations or measurements.

We will briefly present the modified Volterra series equation in section II, then show how this fits efficiently to system level modeling in section III. An application example and possible extension is then presented which shows the effectiveness of the proposed model.

II. CLASSICAL VS DYNAMIC VOLTERRA SERIES EXPANSION

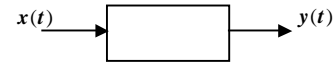


Fig.1 Nonlinear system with memory

The output $y(t_n)$ of a nonlinear system with memory duration $M\Delta t$ can be intuitively expressed as

$$y(t_n) = f(\vec{x}) \quad (1)$$

$$\vec{x} = [x(t_n), x(t_n - \Delta t), \dots, x(t_n - M\Delta t)]^T$$

Taylor series expansion of $y(t_n)$ around some arbitrary signal trajectory $\vec{x} = \vec{x}_0$ gives equation (2) below, from

which we obtain the well known Volterra series expansion (3) as $\Delta t \rightarrow 0$, by choosing $\bar{x}_0 = [0, 0, \dots, 0]^T$

$$y(t_n) = f(\bar{x}_0) + \Delta f(\bar{x}_0)^T (\bar{x} - \bar{x}_0) + \frac{1}{2} (\bar{x} - \bar{x}_0)^T [\Delta^2 f(\bar{x}_0)] (\bar{x} - \bar{x}_0) + \dots \quad (2)$$

$$y(t) = \sum_{n=1}^{\infty} y_n(t) \quad (3)$$

$$y_n(t) = \int_0^{\tau} \dots \int_0^{\tau} \mathbf{h}_n(\lambda_1, \dots, \lambda_n) \prod_{i=1}^n \mathbf{x}(t - \lambda_i) d\lambda_i$$

$\mathbf{h}_n(\lambda_1, \dots, \lambda_n)$ is called Volterra kernel of order n . As it can be seen, Volterra kernels are independent of the input signal $\mathbf{x}(t)$. They are indeed coefficients of a power series expansion. We thus see that for most nonlinear applications encountered in analog IC systems it is necessary to consider a large number of kernels, say up to five and more, in order to accurately describe the response. Here comes the two-fold difficulty of measuring/identifying high order kernels and computing multiple dimensional integrals, which limits usefulness of this elegant approach.

To try resolving these limitations, it was suggested by Asdente et al [10] that carrying Taylor series expansion in eq (2) around a well-behaved trajectory \bar{x}_0 , given by a priori knowledge of the system response, rather than the null vector, can give excellent convergence properties to the resulting series. One simple and efficient trajectory has been found to be [11]:

$$\bar{x}_0 = [\mathbf{x}(t_n), \mathbf{x}(t_n), \dots, \mathbf{x}(t_n)]^T, \quad (4)$$

i.e. the steady state input. Considering this trajectory in (2) yields readily a modified Volterra series of the form below

$$y(t) = y_{dc}(\mathbf{x}(t)) + \int_0^{\tau} \dots \int_0^{\tau} \hat{\mathbf{h}}_n(\mathbf{x}(t), \lambda_1, \dots, \lambda_n) \prod_{i=1}^n [\mathbf{x}(t - \lambda_i) - \mathbf{x}(t)] d\lambda_i \quad (5)$$

In the above expression $y_{dc}(\mathbf{x}(t))$ is the static characteristic (DC) of the system, $\hat{\mathbf{h}}_n(\mathbf{x}(t), \lambda_1, \dots, \lambda_n)$ are the modified Volterra kernels, which we will call hereafter *dynamic Volterra kernels* to imply their dependence on the instantaneous input signal. This modified series has the important property to separate the purely static effects from memory effects, which are intimately mixed in the classical series. Hence if the system is purely static, the series converges only with the static term, irrespective of the input power. The difference term $\mathbf{x}(t - \lambda) - \mathbf{x}(t)$ in the

equation above implies that if the signal period is small compared to the memory duration τ then the series can be truncated to only first order, with good accuracy.

$$\mathbf{y}(t) = y_{dc}(\mathbf{x}(t)) + \int_0^{\tau} \mathbf{h}(\mathbf{x}(t), \lambda) [\mathbf{x}(t - \lambda) - \mathbf{x}(t)] d\lambda \quad (6)$$

A more convenient form of eq (6) is obtained by considering a Fourier integral in the place of convolution to find.

$$\mathbf{y}(t) = y_{dc}(\mathbf{x}(t)) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\mathbf{H}}(\mathbf{x}(t), \omega) \mathbf{X}(\omega) e^{j\omega t} d\omega \quad (7)$$

$$\hat{\mathbf{H}}(\mathbf{x}(t), \omega) = \mathbf{H}(\mathbf{x}(t), \omega) - \mathbf{H}(\mathbf{x}(t), 0)$$

In the above, $\mathbf{H}(\mathbf{x}(t), \omega)$ is actually the small signal transfer function of the system computed around the pump $\mathbf{x}(t)$. Therefore this model does not need complicate measurements or simulation and extraction procedures as would require a high order classical Volterra series expansion. For illustration, if we consider a single transistor modeling, this will require static I/V curves and bias dependent S parameters, which stand respectively for $y_{dc}(\mathbf{x})$ and $\mathbf{H}(\mathbf{x}, \omega)$.

Eq (7) has been with some success to model GaAs FET transistors [10-12]. However, if we consider a complete subsystem like a power amplifier, the assumption for short memory duration with respect to the signal period can no more hold as biasing, filter and matching circuit will cause memory duration much longer than the carrier period. To make this approach possible, it is necessary to carry the subsystem modeling in the complex envelope signal plane, rather than in the real signal space.

III. ENVELOPE DOMAIN MODELING

The basic assumption on system level modeling is that the signal $\mathbf{x}(t)$ at any component port is a band-limited modulation signal on top of a reference carrier frequency as below:

$$\mathbf{x}(t) = \Re e[\hat{\mathbf{X}}(t) e^{j\omega_0 t}] \quad (8)$$

where $\hat{\mathbf{X}}(t)$ and ω_{0x} are respectively the complex envelope (modulation) and reference carrier frequency of the signal $\mathbf{x}(t)$. All idle frequencies are supposed to be sufficiently filtered-out within the subsystem, or constitute otherwise distinct ports.

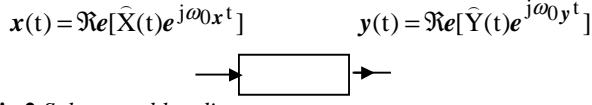


Fig.2 Subsystem bloc diagram

Let us consider the two-port component block depicted in Fig.2 above. The reference carrier frequencies ω_{0x} and ω_{0y} of excitation and response are known a priori. A reference carrier may be a real one or an arbitrary one corresponding to the mid-band frequency of a transmission channel or multiplex.

Now for modeling the block our input/output signals will be $\hat{X}(t)$ and $\hat{Y}(t)$. Because now the memory duration τ of the subsystem is to be compared against the smallest period of the modulation and not the carrier as previously, we may expect the short memory duration assumption to be possible and truncate the Volterra expansion to the first order.

Equation (8) defining an analytic signal [13]. Hence the output envelope $\hat{Y}(t)$ is actually a function of the input envelope $\hat{X}(t)$ and its conjugate. Now reconsidering equation (6), with $\hat{X}(t)$ and $\hat{X}^*(t)$ as input signal and $\hat{Y}(t)$ as output, we readily find

$$\begin{aligned} \hat{Y}(t) = & \hat{Y}_{dc}(\hat{X}(t), \hat{X}^*(t)) + \\ & \int_0^\tau \hat{h}_1(\hat{X}(t), \hat{X}^*(t), \lambda) [(\hat{X}(t-\lambda) - \hat{X}(t))] d\lambda + \\ & \int_0^\tau \hat{h}_2(\hat{X}(t), \hat{X}^*(t), \lambda) [(\hat{X}^*(t-\lambda) - \hat{X}^*(t))] d\lambda \end{aligned} \quad (9)$$

Substituting the convolution integral in (9) by a Fourier integral equivalently yields:

$$\begin{aligned} \hat{Y}(t) = & \hat{Y}_{dc}(\hat{X}(t), \hat{X}^*(t)) + \\ & \frac{1}{2\pi} \int_{-BW/2}^{BW/2} \hat{H}_1(\hat{X}(t), \hat{X}^*(t), \Omega) \times \hat{X}(\Omega) e^{j\Omega t} d\Omega + \\ & \frac{1}{2\pi} \int_{-BW/2}^{BW/2} \hat{H}_2(\hat{X}(t), \hat{X}^*(t), -\Omega) \times \hat{X}^*(\Omega) e^{-j\Omega t} d\Omega \end{aligned} \quad (10)$$

where BW is the signal bandwidth and $\hat{X}(\Omega)$ the input spectrum.

If we consider the particular case of amplifier or frequency converter, we can write:

$$\begin{aligned} \omega_{0y} &= \frac{p}{q} \omega_{0x} \\ p, q &\in Z \\ \omega_{0z} &\in \mathfrak{R} \end{aligned} \quad (12)$$

where ($p=q=1$) for amplifier, ($p>1, q=1$) for multiplier and ($p=\pm 1, q=1$) for mixer.

We then see that because of the time invariance of their input-output relation, the above kernels take the form below, where $|\hat{X}(t)|$ and $\phi_{\hat{X}(t)}$ are the magnitude and phase of $\hat{X}(t)$.

$$\hat{Y}_{dc}(\hat{X}(t), \hat{X}^*(t)) = Y_{dc}(|\hat{X}(t)|) e^{j\frac{p}{q}\phi_{\hat{X}(t)}} \quad (13)$$

$$\hat{H}_1(\hat{X}(t), \hat{X}^*(t), \Omega) = H_1(|\hat{X}(t)|, \Omega) e^{j\frac{p-q}{q}\phi_{\hat{X}(t)}}$$

$$\hat{H}_2(\hat{X}(t), \hat{X}^*(t), \Omega) = H_2(|\hat{X}(t)|, \Omega) e^{j\frac{p+q}{q}\phi_{\hat{X}(t)}}$$

In the above, $Y_{dc}(|\hat{X}(t)|)$ is the static characteristic of the subsystem, i.e. the response of the subsystem under a non modulated carrier excitation $x(t) = \Re e[\hat{X}_0 e^{j\omega_0 x t}]$. This corresponds to the well known AM/AM and AM/PM curve.

$H_1(|\hat{X}(t)|, \Omega)$ and $H_2(|\hat{X}(t)|, \Omega)$ are the *synchronous* and *image* Volterra transfer functions of the subsystem, which the extraction principle is described below.

IV. Extraction of dynamic Volterra series kernels

From Equation (10), we see that it is possible to extract the three kernels by applying at input a two-tone signal of the form

$$x(t) = \Re e[\hat{X}_0 e^{j\omega_0 x t}] + \Re e[\delta \hat{X} e^{j(\omega_0 x + \Omega)t}], \quad |\delta \hat{X}| \ll 1 \quad (14)$$

or in terms of envelope signal, a DC signal plus a small sinusoidal modulation:

$$\hat{X}(t) = |\hat{X}_0| + \delta \hat{X} e^{j\Omega t}, \quad |\delta \hat{X}| \ll 1 \quad (15)$$

In fact, applying (14) into (10), the output signal writes $\hat{Y}(t) = \hat{Y}_0 + \delta \hat{Y}^+ e^{j\Omega t} + \delta \hat{Y}^- e^{-j\Omega t}$, so that:

$$\begin{aligned}
Y_{dc}(\hat{X}_0) &= \hat{Y}_0 \\
H_1(\hat{X}_0, \Omega) &= \frac{\partial \hat{Y}^+}{\partial \hat{X}} - \frac{1}{2} \left[\frac{\partial \hat{Y}_0}{\partial |\hat{X}_0|} + \frac{\hat{Y}_0}{|\hat{X}_0|} \right] \\
H_2(\hat{X}_0, -\Omega) &= \frac{\partial \hat{Y}^-}{\partial \hat{X}^*} - \frac{1}{2} \left[\frac{\partial \hat{Y}_0}{\partial |\hat{X}_0|} - \frac{\hat{Y}_0}{|\hat{X}_0|} \right]
\end{aligned} \tag{16}$$

The extraction procedure is therefore straightforward, as sketched in Fig.3. The three characteristics of the model can be computed by a simple two-tone measurement on a phase calibrated network analyzer [14] or in any steady state simulation tool. The static characteristic $Y_{dc}(\hat{X}_0)$ accounts for the purely static nonlinear effects and the two nonlinear transfer functions $H_1(\hat{X}_0, \Omega)$ and $H_2(\hat{X}_0, -\Omega)$ for the memory effects.

The frequency distance Ω between the two tones is to be swept throughout the bandwidth BW of the subsystem and the input signal magnitude $|\hat{X}_0|$ from linear region to saturation.

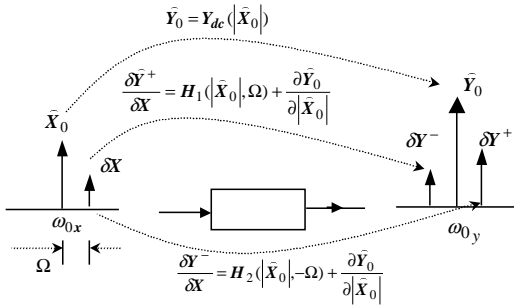


Fig.3 Dynamic Volterra kernels extraction set-up

V. Model implementation

Eq (10) may look a bit cumbersome with the integral notation, actually the model allows a very efficient numerical implementation, by fitting $H_1(\hat{X}(t), \Omega)$ and

$H_2(\hat{X}(t), -\Omega)$ with rational function:

$$H_i(\hat{X}_0, \Omega) = \frac{a_{i0} + \sum_{k=1}^K a_{ik} (\hat{X}_0) \times (j\Omega)^k}{1 + \sum_{p=1}^P b_{ip} (\hat{X}_0) \times (j\Omega)^k} \tag{17}$$

$$i = 1, 2$$

Inserting (17) into (10), we find the expression (18), where we see that the model proposed results in a infinite impulse response filter (IIR) with coefficients depending on the input signal. For many applications, a filter of order 2 to 3

is sufficient. This is therefore suitable to all system-level simulators and can be implemented in analog HDL languages.

$$\begin{aligned}
\hat{Y}(t) &= \tilde{Y}_{dc}(|\hat{X}(t)|) e^{j\frac{P}{Q}\phi_{\hat{X}}(t)} + \hat{Y}_1(t) e^{j\frac{P-Q}{Q}\phi_{\hat{X}}(t)} + \\
&\quad \hat{Y}_2(t) e^{j\frac{P+Q}{Q}\phi_{\hat{X}}(t)} \\
\hat{Y}_1(t) + \sum_{p=1}^P b_{1k}(|\hat{X}(t)|) \frac{d^p \hat{Y}_1(t)}{dt^p} &= a_{10} \hat{X}(t) + \sum_{k=1}^K a_{1k}(|\hat{X}(t)|) \frac{d^k \hat{X}(t)}{dt^k} \\
\hat{Y}_2(t) + \sum_{p=1}^P b_{2k}(|\hat{X}(t)|) \frac{d^p \hat{Y}_2(t)}{dt^p} &= a_{20} \hat{X}^*(t) + \sum_{k=1}^K a_{2k}(|\hat{X}(t)|) \frac{d^k \hat{X}^*(t)}{dt^k}
\end{aligned} \tag{18}$$

VI. MODEL APPLICATION EXTENSION

The experiments we have carried show that the above model is very well behaved up to 1-2 dB compression and more for power amplifiers where the biasing circuit has been carefully taken care to optimize the DC settling time. For illustration, Fig. 4 compares the model prediction against the HB circuit simulation of third order intermodulation ratio (IM3) for a four stage 6 watts MMIC amplifier used in a radar application. The agreement is fairly good between the two. The order of model used was 2 ($P=Q=2$), computation time for the model is only a few seconds, while the harmonic balance simulation takes more than half an hour to complete the power and frequency sweeps. There is some discrepancy between the model and the circuit simulation in the power region where the IM3 exhibits a resonance phenomenon as to the two tones distance. This issue is considered in the model extension suggested below. It is worth noting that computing the IM3 ratio for varying tone distance and power is probably the most effective test for amplifier model, because it highlights the nonlinearity and memory problems, which otherwise is hidden in computing higher level averaging figures like ACPR.

In cases where DC settling time is large due to a poorly designed bias network, AGC loops or severe thermal effects, the memory duration may extend very much and weaken the above assumption of short memory duration. To account for this situation which is common case to narrow band circuits, one may reconsider the expansion in equation (2), replacing the single power basis functions x^n with an arbitrary shaped nonlinear function that achieve a better series convergence. The derivation is then beyond the scope of this paper and may be found in [15]. Doing so, one obtain a slightly modified model equation

below that has proven to be very effective, deep into saturation.

$$\hat{Y}(t) = Y_{dc}(\hat{X}(t), \hat{X}^*(t)) + \int_0^{\tau} h(\hat{X}(t-\lambda), \hat{X}^*(t-\lambda), \lambda) \hat{X}(t-\lambda) d\lambda \quad (19)$$

The kernel of the above equation may be readily extracted by driving the circuit with a unit step input envelope.

Figures 5 and 6 compare the performance of the basic and extended Volterra models against the circuit simulation for a typical narrow band amplifier. One may see a good agreement between the extended model and the circuit simulation, with a substantial improvement of the extended model over the basic modified-Volterra model.

It is worth noting the important variation of the IM3 curves for low tones spacing, even at low input power. This is an important indication that extrapolation of intermodulation distortion from a single tone spacing IP3 can lead to very poor predictions.

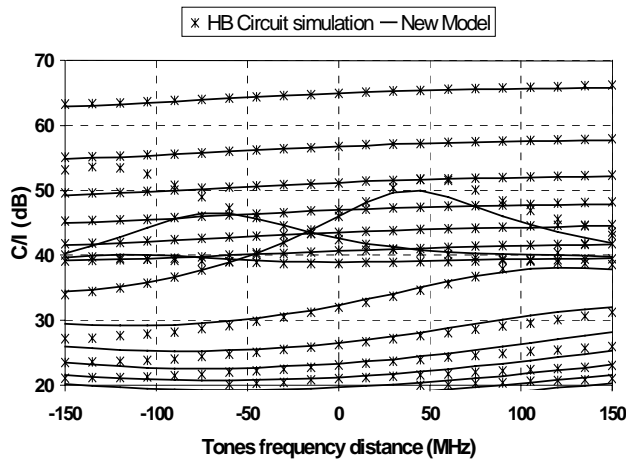


Fig.4 Amplifier IM3 ratio for varying input power (-20dBm to -3dBm) up to 2dB gain compression

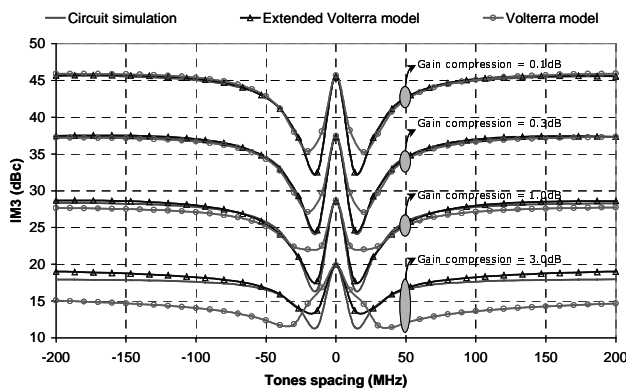


Fig.5 Narrow band amplifier IM3 ratio for varying input power (up to 3dB gain compression) and tone spacing

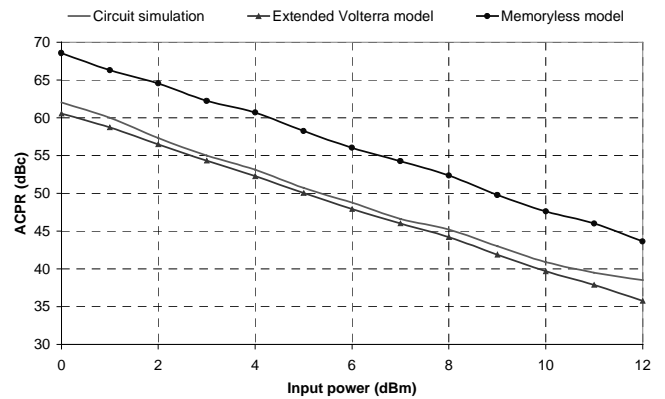


Fig.6 ACPR curve for narrow band amplifier with 20Mb/s qpsk signal

CONCLUSION

The paper has reviewed the basic theory of a powerful behavioral modeling mechanism based on modified Volterra series, so-called dynamic Volterra series model. The model is simple to derive from circuit simulation and also from network analyzer tools. Its implementation has been shown to be suitable to system-level simulators, guarantying simulation speed. It can be applied to various functions like power amplifiers, multipliers and converters to capture nonlinear behavior with memory, especially in modern high speed applications.

REFERENCES

- [1] H.B. Poza et al. "A wideband data link computer simulation model", *Proc. NAECON Conference*, 1975
- [2] A. Saleh, "Frequency-independent and frequency-dependent models for TWT amplifiers", *IEEE Trans. Communications*, COM-29, pp.1715-1720, Nov. 1981
- [3] M. T. Abuelma'atti, "Frequency-dependent Nonlinear Quadrature Model for TWT amplifiers", *IEEE Trans. Communications*, COM-32, pp.982-986, Aug. 1984
- [4] G. Chrisikos, C. J. Clark, A.A. Moulthrop, M.S. Muha and C.P. Silva, "A nonlinear ARMA model for simulating power amplifiers", *IEEE MIT-S Digest*, pp. 733-736, 1998
- [5] W. Bosch and G. Gatti, "Measurement and simulation of memory effects in predistortion linearizers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, No 12, pp. 1885-1890, Dec. 1989.

- [6] J. Roychowdhury, "Reduced order modeling of time varying systems", *IEEE Trans. Circuits and Systems – II: Analog and Digital Signal Processing.*, vol. 46, No 10, pp. 1273-1288, Oct. 1999
- [7] E. Gad, R. Khazaka, M.S. Nakhla and R. Griffith, "A circuit reduction technique for finding the steady state solution of nonlinear circuits", *IEEE Trans. Microwave Theory and Techniques.*, vol. 48, No 12, pp. 2389-2396, Dec. 2000.
- [8] J.R. Phillips, "Projection based approaches for model reduction of weakly nonlinear, time varying systems", *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, No 2, pp. 171-187, Feb. 2003
- [9] M. Rewienski and J. White, "A trajectory piecewise linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices", *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, No 2, pp. 155-170, Feb. 2003.
- [10] M. Schetzen, *The Volterra and Wiener theories of nonlinear systems*, Wiley, 1980
- [11] M. Asdente, M.C. Pascucci and A.M. Ricca, "Modified Volterra-Wiener functional method for highly nonlinear systems", *ALTA FREQUENZA*, vol. XLV, pp. 756-312 E-315 E-759, Dec 1976
- [12] F. Filicori and G. Vanini, "Mathematical Approach to Large Signal Modelling of Electron Devices", *Electronics Letters*, vol.27, pp.357-359, Feb. 1991
- [13] S. Haykin, *Communication Systems*, John Wiley & Sons, New York, 1978
- [14] N. Legallou et al, " A novel measurement method for the extraction of dynamic volterra kernels of microwave power amplifiers" submitted to GAAS 2000, Paris, 2000
- [15] A. Soury et al, " Measurement based modelling of power amplifiers for reliable design of modern communication systems" ", *IEEE MTT-S Digest*, pp. 795-798, 2003