Separation of the Nonlinear Source–Pull From the Nonlinear System Behavior

Philippe Crama, *Member, IEEE*, Yves Rolain, *Senior Member, IEEE*, Wendy Van Moer, *Associate Member, IEEE*, and Johan Schoukens, *Fellow, IEEE*

***Abstract*—In this paper, we describe a measurement technique to identify a nonlinear system in the presence of nonlinear source–pull. Models identified with continuous-wave measurement data are not generalizable when nonlinear source–pull is present. This is demonstrated on measured data and compared with the performance of the proposed technique. The method is based on two-tone signals with very close frequencies that excite the system at the fundamental and harmonic frequencies. By varying the phase relation between the beat components, the system's nonlinear behavior is separated from the nonlinear source–pull. Note that such excitations can be generated using commercially available synthesizers with single-sidebnad in-phase-quadrature modulation options.**

***Index Terms*—Nonlinear distortions, nonlinear source–pull.**

I. Introduction

In recent years, a growing interest appeared for modeling the nonlinear behavior of microwave circuits. Prototypes of nonlinear vectorial network analyzers became available (see [1] and [5]), and the first attempts were made to model the measured nonlinear behavior using Volterra series approaches (see [2] and [4]). In many cases, the aim is to model the relation between the fundamental incident wave $A_1(f)$ and the nonlinear distorted output waves $B_2(f)$, $B_2(2f)$, and $B_2(3f)$. Although this seems at first glance to be a simple task, it turns out that the interaction between the nonlinear circuit and the continuous wave (CW) generator (nonlinear source–pull) complicates this picture significantly. Due to these effects, the system is not only excited by the pure carrier $A_1(f)$ alone, but also by its higher harmonic components $A_1(2f)$ and $A_1(3f)$.

In principle, it should be possible to measure and identify the full Volterra map that describes the relation between $A_1(f)$, $A_1(2f)$, $A_1(3f)$ and $B_2(f)$, $B_2(2f)$, $B_2(3f)$. However, the experiment is not "rich" enough to get a reliable estimate. The user sets only $A_1(f)$ on the CW source: $A_1(2f)$ and $A_1(3f)$ are unwanted contributions created by the source–device-under-test (DUT) pair. As these components are determined by the setup, they do not vary independently of $A_1(f)$. This leads to an excitation that does not fill up the full five-dimensional excitation space, but remains instead on a one-dimensional curve in it. The five dimensions are the amplitude of each component (three in total) and the phase of the harmonics referred to the fundamental (two more dimensions).

Manuscript received March 17, 2001. This work was supported by Funds for Scientific Research Flanders-Vlaanderen, by the Flemish community under Concerted Action Interpretation by Measurement, Modeling, and Identification, and by the Belgian Government under InterUniversitaire AttracticPool-4/2.

The authors are with the Electrical Measurement Department, Vrije Universiteit Brussel, B-1050 Brussels, Belgium (e-mail: Philippe.Crama@vub.ac.be).

Publisher Item Identifier 10.1109/TMTT.2002.801318.

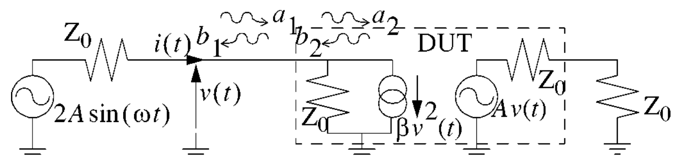


Fig. 1. Ideal source driving an amplifier (the DUT) with a nonlinear input impedance.

When a model is extracted using this type of measurement, a good agreement between model and measurement is obtained. However, the resulting model is very unreliable, as it does not allow any generalization. Even simple simulations like the prediction of the system output for a pure sine excitation at the input (putting $A_1(2f)$ and $A_1(3f)$ equal to zero) will fail. The extracted model uses the additional degrees of freedom contained in the input space to reach a better fit of the measured data: basically, the model does not only describe the nonlinear system behavior, but the combination of the generator and system setup. Changing one of the components of this setup will, hence, change the model. It is clear that, for simulation purposes, this is unacceptable. For that reason, the nonlinear source–pull should be separated from the system model. In this paper, the quality of a model will be judged by the plausibility of its response for an ideal sine wave at the input.

II. Illustration of the Problem

Nonlinear source load–pull is an effect induced by the cascade of a source with nonzero output impedance and a nonlinear device. Even for a generator that is matched for all frequencies and whose output spectrum is a pure tone, there will still be waves at the harmonic frequencies exciting the DUT. Consider the following system built with a perfect generator driving an amplifier with a nonlinear input impedance. We can see that even though the generator is supposed to be matched perfectly, harmonics appear in the excitation spectrum because of the variation of the input impedance of the DUT implied by the nonlinearity itself.

Using the circuit of Fig. 1, the incident (a_1) and reflected (b_1) wave are easily calculated if the square root is approximated by its Taylor series

$$a_1 = A \cdot \sin(\omega t)$$
$$b_1 = \frac{1}{2}\beta Z_0 A^2 \sin^2(\omega t) + \frac{1}{2}\beta^2 Z_0^2 A^3 \sin^3(\omega t). \quad (1)$$

However, due to the nonlinear input impedance, the characteristic impedance of the input of the DUT changes ($\neq Z_0$) and now depends on the power in a_1, resulting in a power-dependent mismatch at the device input. Converting the waves to another impedance system Z_{NL} will introduce the harmonics in the wave traveling forward; (2) shows that the DUT will be excited by the nonlinear harmonics, even though our generator is supposed to be perfect. This shows that when looking at the nonlinear behavior of a DUT, one has to take into account that the DUT is always excited at the harmonic frequencies.

Even if a filter would be added in between the generator and DUT, this would not solve the problem, as the impedance Z_0 would now be the output impedance of the filter. Note also that this mismatch is dependent on the input power and frequency; a linear matching network, albeit ideal, can, therefore, never be used to obtain a perfect match for all cases.

Hence, the nonlinear source–pull effect cannot be avoided for devices whose input impedance is nonlinear. The goal of this paper will now be to avoid that this disturbance impairs the quality of the models that are extracted. If no specific precaution is taken, the extracted model will describe the generator–DUT pair as a whole. Modifying one of the components (e.g., using a spectrally pure excitation instead of the actual distorted signal) may—and actually will—significantly degrade model performance

$$
\begin{aligned}
a_2 = {} & \frac{Z_0 + Z_{NL}}{2Z_0} A \cdot \sin(\omega t) + \frac{Z_0 - Z_{NL}}{2Z_0} \\
& \cdot \left(-\frac{1}{2} \beta Z_0 A^2 \sin^2(\omega t) + \frac{1}{2} \beta^2 Z_0^2 A^3 \sin^3(\omega t) \right) \\
b_2 = {} & \frac{Z_0 - Z_{NL}}{2Z_0} A \cdot \sin(\omega t) + \frac{Z_0 + Z_{NL}}{2Z_0} \\
& \cdot \left(-\frac{1}{2} \beta Z_0 A^2 \sin^2(\omega t) + \frac{1}{2} \beta^2 Z_0^2 A^3 \sin^3(\omega t) \right). \qquad (2)
\end{aligned}
$$

III. Basic Idea

From the identification point-of-view, the main trouble comes from the fixed phase relation between the different harmonics of the input signal. One could use an impedance tuner between the source and DUT to vary the source–pull and change the phase relation between the harmonics. However, this does not allow to control directly the phase of the harmonics unless complex harmonic tuners are used.

The solution that we propose in this paper is to excite the system with a two-tone excitation with almost coinciding frequencies f_1 and f_2. This is very close to a CW excitation, but introduces one more user controllable degree of freedom in the choice of the input signal: the phase between the two components. Now, when mixing will occur between these components, the resulting phase will be influenced by the phase difference between the carriers at f_1 and f_2.

Due to the nonlinear source–pull, the incident waves will consist of $A_1(f_1)$, $A_1(f_2)$ and their (inter)modulation products. All these components contribute to the output. Restricting ourselves without loss of generality to third-degree nonlinearities, the contributions at the fundamental frequency f_1 have the following form: $A_1(f_1)$ (linear), $A_1(f_1)A_1(3f_1)A_1(-3f_1)$, and other third-degree participations. Similar contributions appear at f_2: $A_1(f_2)$ (linear), $A_1(f_2)A_1(3f_2)A_1(-3f_2)$ together with other third-degree participations. Completely different terms appear at the nearby frequency $3f_1 - 2f_2$, e.g., $A_1(f_2)A_1(3f_1)A_1(-3f_2)$. The Volterra kernels [3] associated with these different third-order products can be approximated by the same value because these contributions hit the nonlinear system at almost the same place on the multidimensional frequency grid ($f_1 \approx f_2$). However, they have a completely different phase relation that is easily changed by varying the phase of $A_1(f_1)$ and $A_1(f_2)$ using the single-sideband (SSB) in-phase quadrature (IQ) modulation.

Fig. 2. Experimental setup.

Theoretically speaking, any system that can be modeled using Volterra series can be identified using this method. However, for strong nonlinearities, the number of harmonics to take into account might make the computation cost prohibitive: the more harmonics are generated, the more input products one has to compute.

IV. Experimental Verification

The prediction power of the extracted model will be assessed as follows: the model that is extracted using the proposed dual-tone method is used to model the DUT's response to a single sine wave at the same frequency. Next, the DUT's response itself is measured at this frequency for a sine-wave experiment and compared to the obtained model and output power predicted by the model extracted from the single-tone measurements.

A. Measurement

The proposed method was tested on a MAR 6 amplifier from Mini-Circuits, Brroklyn, NY, with the 1-dB compression point at −6-dBm input power at 900 MHz. The measurements were made on the same system (Fig. 2) with two different excitation signals: a single tone at 900 MHz and a dual-tone excitation at 899.995 and 900.005 MHz. This allows to compare how both methods perform. The input and output waves were measured for an input power going from −20 to 0 dBm in 1-dB steps. For each wave, the necessary frequency components were recorded with the vectorial nonlinear network analyzer described in [1].

Fig. 3 shows that the phase differences between the fundamental and second harmonic (1800 MHz) are spread over the whole range $[-\pi; \pi]$ for the dual-tone experiments as intended, while the phases of the second component for the single-tone experiments stay the same. The plotted phase is computed as the phase difference between the component at 1800 MHz and the square of the first component (900 MHz for the single-tone measurements, 899.995 MHz for the dual-tone measurements). This representation cancels out the influence of the phase of the fundamental and allows to see clearly the influence of the eight phase steps used in the excitation signal. For low-input

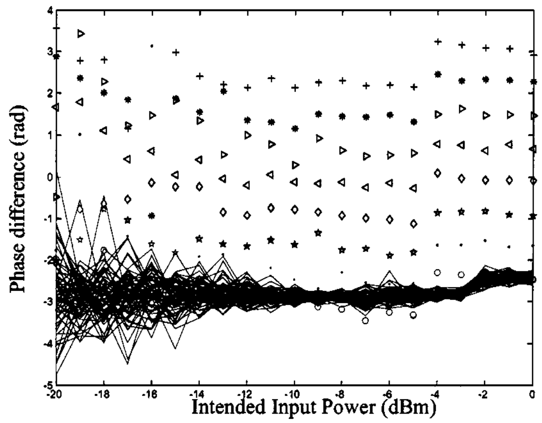


Fig. 3. Comparison of the phase of the second input harmonic for the MAR 6 setup [—: for single-tone measurements, symbols: for the dual-tone experiment (eight different phases/experiments)].

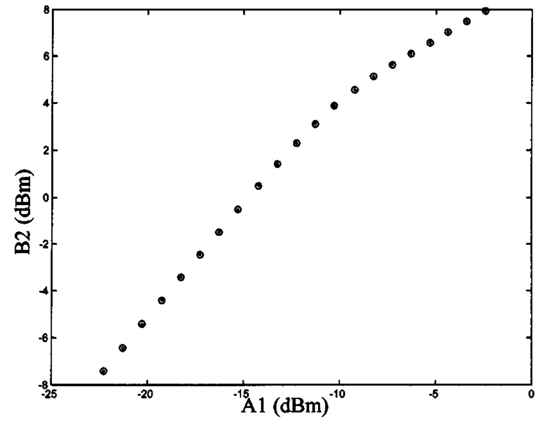


Fig. 4. Comparison of the measured output power and the fitted output power for the single-tone experiments for the MAR 6 (○: fitted data, •: measurements).

powers, the second harmonics are very close to the noise floor, resulting in a random distribution of the phases. For the higher input powers, the second harmonics are clearly defined and it is obvious that the phase relation between the second harmonic and fundamental stays fixed and that the variation comes from the different phase realizations of the dual-tone excitation.

B. Extracted Models

The kernels were estimated with (3) as the cost function. This is a slightly modified weighted least squares estimator [7] where $B_2(i)$ is the measured output wave, $X(i)$ is a horizontal vector containing the products of the measured input waves (A_1), V is a vertical vector with the unknown Volterra kernels (ordered in the same way as the products), and w_i is the weighting function

$$\mathrm{CF} = \sum_i \frac{\left\| \log \dfrac{B_2(i)}{X(i) \cdot V} \right\|^2}{w_i^2}. \tag{3}$$

In identification, it is common practice to select w_i to match the level of uncertainty: w_i is then equal to the standard deviation of the numerator of (3). Here, the weights are not computed analytically using the standard deviation of the measurements, but estimated from the measured values: each measurement has been repeated (64 times for the CW measurements and three times for the dual-tone measurements), and the standard deviation of the equation error is computed using each repetition as a sample. This makes the weights a function of the estimated kernels. To simplify the equation, the kernel values from the previous estimation step are used. Since the proposed scheme is iterative, the estimation of the kernel values has to be initialized. The first estimates of the kernels are the values that minimize the classical unweighted least squares cost function.

The measurements are done over a wide power range. To cover the whole range, attenuators have been added during the experiment to optimize the dynamic range of the analog-to-digital convertors (ADCs). This causes the noise level in the measurements to increase with the input power level.

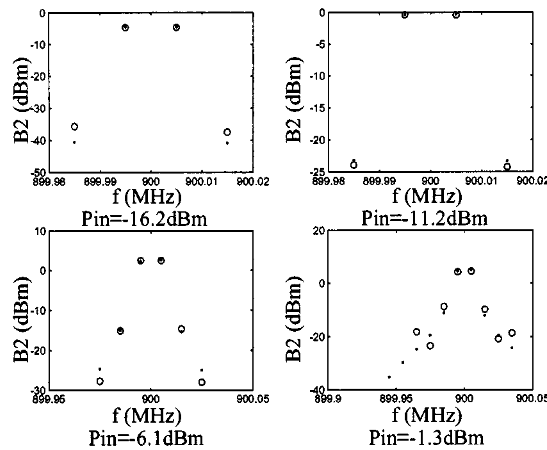


Fig. 5. Quality of the fit for some of the MAR 6 amplifier's dual-tone measurements ($\Delta\phi_{f1,\,f2} = \pi$, $\cdot$: measurements, ○: model value).

C. Comparison of the Extracted Models for the Single- and Dual-Tone Measurements

With the data from the CW experiment, the Volterra kernels up to the seventh order contributing to the fundamental frequency at the output were estimated. These kernels will be called the single-tone kernels, with "single tone" indicating they were obtained from CW measurements. In Fig. 4, the model was applied on the same experimental data that is used to identify the model and compared to the measured output. A perfect fit is obtained. The convergence of the kernel values is extremely fast: the optimum is reached usually in less than five iterations.

It is impossible to represent on a plot the different axes needed to compare the fit of the dual-tone model with the dual-tone measurements: there should be an axis for the input power, one for the phase realization, one for the frequency, and one for the output power. Thus, only a subset of the data is plotted in Fig. 5 to illustrate the comments. For high-input powers, additional frequency lines close to the fundamental frequency are included: these are the lines excited by the nonlinear source–pull that grow above the noise when the input power increases. Their existence in the bottom two plots of Fig. 5 proves that nonlinear source–pull is indeed present.

The fit of the dual-tone model is not as good as for the single-tone case: the output predicted by the model for the

TABLE I
NUMERICAL VALUES OF SOME OF THE ESTIMATED VOLTERRA KERNELS FOR THE MAR 6 AMPLIFIER

Volterra kernel	Single Tone kernel value	CW Power	Dual Tone kernel value	Dual Tone Power
$V(f_1)$	-0.2096387-5.565479i	12.5dBm	-2.955302+4.868292i	13.82dBm
$V(f_1,f_1,-f_1)$	-38.33139+2.631012i	1.9dBm	+13.32411-117.6923i	19.0dBm
$V(f_1,f_1,f_1,-f_1,-f_1)$	+8867.484+4979.624i	23.9dBm	-53.37062+1165.927i	19.6dBm
$V(f_1,f_1,f_1,f_1,-f_1,-f_1,-f_1)$	+122711.0+5927.236i	19.5dBm	+649.5390-2447.282i	7.8dBm
$V(2f_1,-f_1)$	-499.2629-61.82980i	1.6dBm	+309.8736+69.61719i	-2.1dBm
$V(3f_1,-f_1,-f_1)$	-12208.92+15037.13i	-7.4dBm	+1188.959-765.3288i	-11.1dBm
$V(f_1,f_1,f_1,-2f_1)$	+71593.87-18381.58i	14.1dBm	+3886.299-5189.947i	0.4dBm
$V(2f_1,f_1,-f_1,-f_1)$	+7312.453+22897.17i	13.8dBm	-8263.892-2.454021i	12.1dBm
$V(3f_1,f_1,-f_1,-f_1,-f_1)$	+320332.0-409185.2i	0.7dBm	-3402.645+1220.194i	-15.9dBm
$V(f_1,f_1,f_1,f_1,-f_1,-2f_1)$	-1381997+1937081i	24.8dBm	-23616.06+20997.73i	2.4dBm
$V(2f_1,f_1,f_1,-f_1,-f_1,-f_1)$	+507612.4-1422140i	26.9dBm	+46275.19-7148.897i	11.8dBm

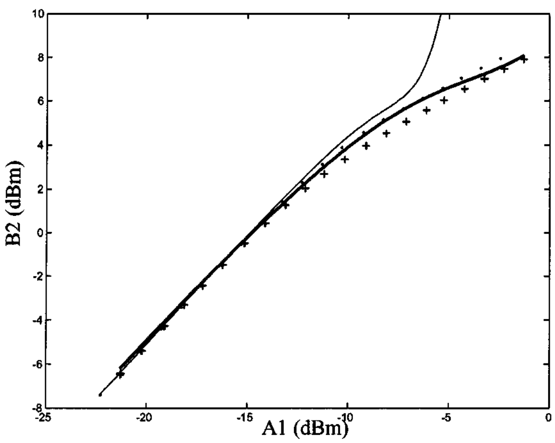


Fig. 6. Output of the identified model for a pure sine excitation for the MAR 6 (—: output of single-tone model, ■: output of dual-tone model, ·: single-tone measurements, +: dual-tone measurements).

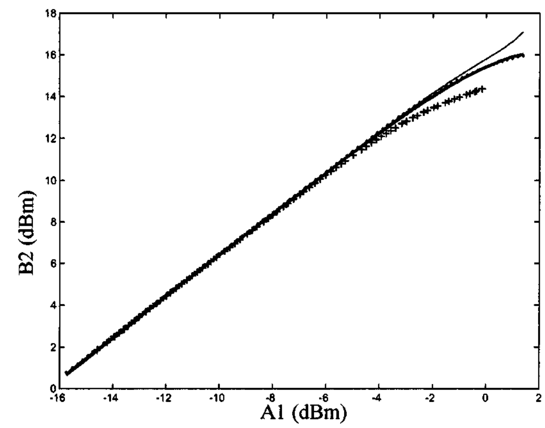


Fig. 7. Output of the identified model for a pure sine excitation for the RFIC (—: output of single-tone model, ■: output of dual-tone model, · : single-tone measurements, +: dual-tone measurements).

measured input matches the measured outputs less closely. This was to be expected: the input signal is much richer, yet the same number of parameters is estimated. Note also that the measurement points where the fit appears to be much worse correspond to frequency lines that are not in the intended excitation spectrum. These lines result of the nonlinear source–pull third-order contributions (e.g., $2f_1 - f_2, 2f_2 - f_1$). These contributions are much smaller than the lines excited by the generator and have a bad signal-to-noise ratio. Since the estimator weights measurements according to their quality, the fit for these lines may be loose without a significant increase of the cost. For lower input powers (see the upper plots of Fig. 5, where these side lines are hidden in the noise), they are not even taken into account.

A simulation of the CW and dual-tone model is performed in Fig. 6 for a pure sine-wave excitation using the models extracted from the measurements, using only the kernels $V(f_1)$, $V(f_1, f_1, -f_1)$, $V(f_1, f_1, f_1, -f_1, -f_1)$, and $V(f_1, f_1, f_1, f_1, -f_1, -f_1, -f_1)$. Despite the perfect fit between the CW model and measurement (all input harmonics included), a very poor prediction is obtained here, while the dual-tone model predicts much better.

A second experiment was performed on an RF amplifier from Motorola, Denver, CO, i.e., the MRFIC2006. The measurements were done at 600 MHz, from -16-dBm input power until 2-dB compression. The 1-dB compression point of this amplifier is at -0.2-dBm input power at 600 MHz. The results for this second experiment were very analogous. Thus, only one figure (Fig. 7) assessing the prediction power of the extracted model will be included.

Fig. 7 shows the same simulation as Fig. 6 for the RF integrated circuit (RFIC). The model extracted from dual-tone measurements was extrapolated to match the CW power range and still looks plausible: the crosses indicate that the dual-tone measurements cover a smaller power range than the thick line, which shows the output of the dual-tone model. On the other hand, the output power simulated with the model identified from the single-tone model does not look plausible at all since expansion is predicted instead of the measured compression.

D. Comparison of the Estimated Kernel Values for the MAR 6

Table I shows the estimated single- and dual-tone kernel values for the MAR 6 DUT. To compare their relative importance, the power of each kernel for the highest measured input power of -2 dBm is included. This power is computed as the predicted output power when all other kernel values are set to zero. The bold kernels in Table I are the ones used for the predictions in Fig. 6.

The first line shows that we are in compression: the output power predicted by the linear kernel is higher than the measured output power. The higher order kernels will, hence, need to cancel a part of this linear contribution. Since these kernels tend to infinity for rising power, a saturation contribution as required here can only be obtained when kernel contributions are almost canceling. This explains the huge power values created by separate kernels, for instance, the $V(f_1, f_1, -f_1)$ kernel. However, for the higher order kernels, the contributions of the single-tone kernels to the prediction are higher than the contributions of the dual-tone kernels. Due to the phase variation, fitting the dual-tone kernels to reach cancellation on a certain phase realization will lead to a huge error for another phase realization of the input signal at the same input power. Hence, the estimation algorithm keeps the power from these kernel values low instead of trying to follow the noise. In the CW case, the estimator is able to follow the noise by increasing the higher order kernel's power as long as they nearly cancel out because the excitation space is not completely filled by the input signal.

Note that for the MAR 6, a good estimation could also have been obtained from single-tone measurements. One could simply consider the source as being perfectly matched and identify the Volterra kernels linked with the fundamental only. However, this works only because the matching of the generator and the MAR 6 was so good that the source–pull could hardly be seen. For a larger nonlinear source–pull, it becomes mandatory to include the input harmonics, leading to the aforementioned problems. The advantage of the dual-tone experiments is to provide a robust method that can be used without much prior knowledge (such as knowing whether there is nonlinear source–pull or not).

V. Conclusion

A simple method has been proposed to measure and identify sensible models of the nonlinear system behavior, even in the presence of nonlinear source–pull, using a single generator with SSB IQ modulation options (e.g., the Rohde and Schwarz SMIQ 06B generator). As opposed to models identified from CW experiments, these models allow generalization and can be used for simulation purposes. This method has been applied successfully to experimental measured data.

References

[1] T. Van den Broeck and J. Verspecht, "Calibrated vectorial nonlinear-network analyzers," in *IEEE MTT-S Int. Microwave Symp. Dig.*, San Diego, May 23–27, 1994, pp. 1069–1072.

[2] F. Verbeyst and M. Vanden Bossche, "VIOMAP: The S-parameters equivalent for weakly nonlinear RF and microwave devices," *Trans. Microwave Theory Tech.*, vol. 42, pp. 2531–2535, Dec. 1994.

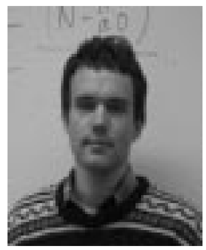
[3] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley, 1989.

[4] Y. Rolain, P. Vael, W. Van Moer, and A. Barel, "Complete 3-port measurement of microwave mixers using a nonlinear vectorial network analyzer," in *IEEE MTT-S Int. Microwave Symp. Dig.*, vol. 3, Baltimore, MD, June 7–12, 1998, pp. 1491–1494.

[5] W. Van Moer and Y. Rolain, "Calibration of a wideband IF nonlinear vectorial network analyzer," in *53rd ARFTG Conf. Dig.*, Anaheim, CA, June 18, 1999, pp. 98–103.

[6] W. Van Moer, Y. Rolain, and J. Schoukens, "An automatic harmonic selection scheme for measurements and calibration with the nonlinear vectorial network analyzer," in *IEEE MTT-S Int. Microwave Symp. Dig.*, vol. III, Boston, MA, June 11–16, 2000, pp. 1855–1858.

[7] W. Van Moer, "New measuring and modeling techniques for RF IC's and their nonlinear behavior," Dept. Appl. Sci., Ph.D. dissertation, Vrije Univ. Brussel, Brussels, Belgium, 2001.



Philippe Crama (M'01) was born in Belgium in 1977. He received the Engineer degree from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 2000.

He is currently an Aspirant of the Fund for Scientific Research Flanders [Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO)]. His primary research concerns the field of system identification for nonlinear systems.

Yves Rolain (M'90–SM'96) received the Burgerlijk Ingenieur degree, Masters degree in computer sciences, and Doctor degree in applied sciences from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 1984, 1986, and 1993, respectively.

He is currently a Senior Research Assistant with the Electrical Measurement Department (ELEC), VUB. His main research interests are microwave measurements and modeling, applied digital signal processing, and parameter estimation/system identification.

Wendy Van Moer (S'98–A'01) received the Electrotechnical Engineer degree in telecommunication and Doctor degree in applied sciences from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 1997 and 2001, respectively.

She is currently a Researcher of the Instituut voor de aanmoediging van innovatie door Wetenschap en Technologie In Vlaanderen (IWT), Electrical Measurement Department (ELEC), VUB. Her main research interests are nonlinear microwave measurement and modeling techniques.

Johan Schoukens (M'90–SM'92–F'97) was born in Belgium in 1957. He received the Engineer degree and Doctor degree in applied sciences from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 1980 and 1985, respectively.

In 1992, he became member of de Coppighaerdt. He is currently a Research Director with the National Fund for Scientific Research (NFWO) and a Part-Time Lecturer with the VUB. His major interests are in the field of system identification for linear and nonlinear systems.