Practical Issues on RF Modelling of Multi-rate Nonlinear Systems

Telmo Reis Cunha, José Carlos Pedro

Resumo – **Este artigo apresenta uma análise sobre questões práticas referentes à utilização de modelos RF de sistemas não lineares que possuem fenómenos em bandas de frequência muito distintas (denominados por sistemas multi-ritmo). Esta característica dificulta a utilização de modelos que transportam para a banda-base o comportamento do sistema, sendo aplicados os chamados modelos RF (que não consideram qualquer translação em frequência das suas características).**

Havendo um interesse cada vez maior na modelação de sistemas multi-ritmo, por exemplo, para analisar fenómenos térmicos em equipamento RF, este artigo aponta alguns problemas de ordem prática na utilização de modelos RF, nomeadamente a enorme quantidade de parâmetros necessários para o modelo, e os elevados tempos de computação associados. São também propostas algumas linhas de acção para contornar estes problemas.

Abstract – **This paper presents an analysis on practical issues regarding the use of RF modelling of multi-rate nonlinear systems. Multi-rate systems are characterized by a frequency spectrum concentrated on distinct and separated frequency bands. This characteristic makes impossible the use of base-band models (that consider frequency shifting of the system spectrum), and so, RF modelling is naturally applied.**

Noticing the growing interest on multi-rate system modelling, for example, to analyse temperature phenomena on RF equipment, this paper presents some implementation problems associated to the use of multi-rate RF models, namely the huge amount of model parameters, and the associated computation time. Some suggestions to work around these problems are also presented.

1. INTRODUCTION

The telecommunication society shows, throughout the last decades, a growing interest on nonlinear behavioural system modelling, as can be verified by the amount of published papers on the issue ([1] through [17] is just a small sample).

The predominant mathematical strategies used in nonlinear system modelling are: neural networks and multidimensional polynomial series (Volterra series). In this paper only the Volterra series approach will be considered, although the conclusions obtained from the presented analysis are extendable to the neural network case. Information on neural network modelling can be found in diverse documentation, in which [18] to [24] are some reference examples.

Volterra series are used in systems that are only mildly nonlinear, that is, whose nonlinearities can be approximated by a low order polynomial expansion around some quiescent point [8], [10], [11]. Fortunately, since nonlinearity is most of the times an undesirable source of signal fidelity impairments, many systems encountered in telecommunications and instrumentation fit that description. Equation (1) shows the Volterra series expansion, in its discrete form, of the transfer function of the system presented in Fig. 1. Notice that the Volterra series is nothing more than the sum of multidimensional convolutions between the input signal and the multidimensional impulse responses (the various n^{th}-order kernels).

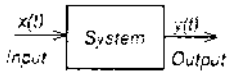

Fig. 1 - General system.

$$y(t) = \sum_{n=0}^{\infty} y_n(t) \qquad where$$
$$y_n(t) = \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_n=0}^{\infty} h_n(\tau_1, \cdots, \tau_n)\, x(t-\tau_1) \cdots x(t-\tau_n) \tag{1}$$

Both representations – neural networks and Volterra series – have advantages and disadvantages. For example, the parameters of a neural network result from a training procedure performed over the application of a known excitation sequence to its input. Different parameter sets can result if different training sequences are used, so their predictive ability can be compromised if the model input has some characteristics that fall outside the expected behaviour they were supposed to have. Volterra series, similarly to one-dimensional polynomials, have severe convergence problems when the input exceeds a certain interval around the quiescent point.

It is not the objective of this paper to analyse theoretical limitations of such methods, being the reader suggested to [8], [10], [11] for further details on the subject. This article is more concerned with practical implementation aspects of nonlinear FIR filters when multi-rate behaviour is to be modelled.

Taking as an example a power amplifier working at some RF frequency band, its behaviour is likely to change if the temperature of its components varies. In many cases, such behavioural changes are quite noticeable, being worthy to model. Since temperature variation is clearly a very low

frequency phenomena, when compared to the RF working band of the amplifier, this becomes a multi-rate system. Evidently, there are many other equipment and phenomena that present this multi-rate characteristic.

If multi-rate information is to be considered by the model, it makes no sense to use a low-pass equivalent model, otherwise part of the information would be lost. So, the solution adopted by several authors is to model the system without any frequency shifting scheme, resulting in the so called RF model.

Since the objective of these models is to simulate the system behaviour to a broad class of inputs, the parameters required for its characterization are the time-domain parameters: the impulse response (first-order kernel) in the case of a linear system, or the multidimensional kernels of the Volterra series (1) in the nonlinear case.

This paper shows that the number of parameters of the general time-domain Volterra series of a multi-rate system is huge, which can compromise the practical use of such technique if some restricting strategies are not applied. A large set of parameters means, on one hand, that the system identification process is very hard and, on the other hand, the computation time during simulation can rapidly become unbearable.

The following Section demonstrates that, for a multi-rate linear system (and also for the nonlinear case), the respective impulse response has a higher set of parameters than the frequency-domain representation, although both contain the same information. Some system examples are given to illustrate the inefficiency of these models.

The third Section of this paper suggests a strategy to avoid having such a huge number of parameters to deal with. This is achieved by restricting the space of non-zero parameters, using a predefined model topology. An example with first and third-order kernels will be thoroughly analysed.

II. TIME DOMAIN REPRESENTATION OF MULTI-RATE SYSTEMS

It is known that the time domain representation of a linear system is the system impulse response, which is the inverse Fourier transform of the frequency-domain representation. In the present case, the Discrete Fourier Transform (DFT) will be considered.

Let us analyse the linear system whose frequency-domain representation is given in Fig. 2 (only the amplitude is depicted). Applying the inverse DFT results in the system impulse response shown in Fig. 3.

From the Fourier transform properties, the following remarks are taken:

- the sampling period Δt of $h(t)$ is the inverse of twice the maximum frequency of $H(j\omega)$;
- the period T of $h(t)$ is the inverse of Δf, which is equivalent to saying that T can be small if $H(j\omega)$ is smooth.

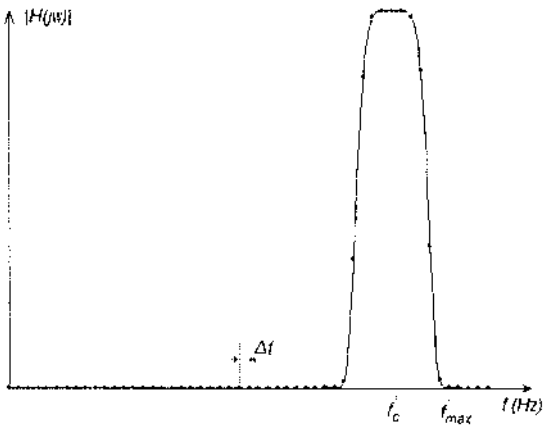

Fig. 2 - Frequency domain representation of a linear system.

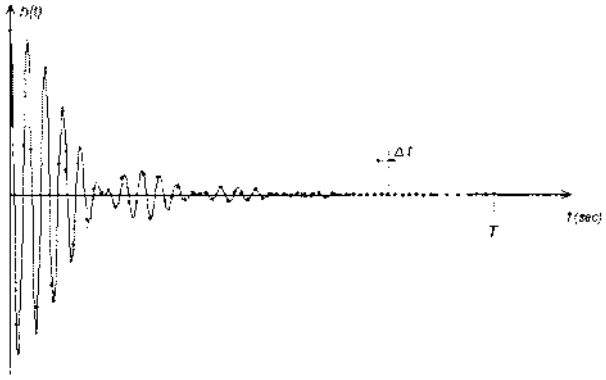

Fig. 3 - Time domain representation of the system of fig.2.

What is immediately observed in Fig. 2 and Fig. 3 is that the frequency-domain representation requires only the knowledge of 11 non-zero parameters while the time domain representation requires at least 36! But both should represent the same information, so why is the number of non-zero parameters so different?

Let us look at the information that each representation contains. As is known, the envelope of $h(t)$ is nothing but the inverse DFT of $H(j\omega)$ when it is shifted to base-band (through a band-pass sampling process [25], [26], for example). Then, $h(t)$ can be formed by filling the inside of the envelope with a cosine wave at the centre frequency f_c of $H(j\omega)$ (obviously, the cosine wave is multiplied by the envelope), and sampled with period Δt.

So, the information of each representation can be described by:

Frequency Domain:

- Curve parameterization of $H(j\omega)$ (by means of a polynomial, for example) → n parameters;
- Centre frequency f_c → 1 parameter;
- Bandwidth Bw → 1 parameter;
- Frequency spacing Δf → 1 parameter;
- All values outside $[f_c - B/2, f_c + B/2]$ are zero (this statement is information too);
- $H(j\omega)$ has even symmetry on the amplitude and odd symmetry on the phase.

Time Domain:

- Curve parameterization of the envelope of $h(t)$ (by means of a polynomial, for example) → n parameters;
- Frequency f_c of the cosine wave multiplying the envelope → 1 parameter;
- Sampling period Δt → 1 parameter;
- Period T of $h(t)$ → 1 parameter;
- $h(t)$ is real and has even symmetry.

The similarities are evident, and the only difference is that $h(t)$ has no statement saying that a certain number of its values are zero! This means that the samples $h(t_k)$ can be non-zero, for all t_k in the period T!

This result, as can be verified through equation (2), is an evident consequence of the definition of the inverse DFT.

$$h(t_k) = \sum_{i=-\infty}^{\infty} H(\omega_i)\, e^{j\omega_i t_k} \qquad (2)$$

On systems like the one presented in Fig. 2, or even in systems having fundamental and harmonic bands, some schemes are usually implemented to avoid handling of a huge number of time domain parameters. As shown in Fig. 4, a sub-sampling (a procedure called band-pass sampling) followed by a proper low pass filtering can concentrate all the non-zero information of $H(j\omega)$ at low frequencies, which naturally produces an impulse response with fewer parameters than the original $H(j\omega)$ (Δt is wider and T remains the same).

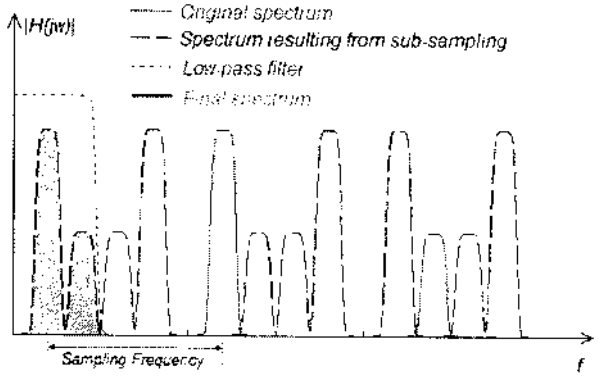

Fig. 4 - Band-pass sampling example, followed by low-pass filtering.

If the system has non-negligible multi-rate characteristics where the location of the distinct bands does not follow a general relationship (such as with the harmonic bands), it can be complicated to apply a frequency down-shifting scheme. Again, if multi-rate characteristics are to be preserved by the model, the low-pass equivalent model is not a valid option. A solution adopted by several authors is simply to use the RF model, where no frequency shift is considered. A typical situation is in the analysis of temperature effects (very low frequency) on systems that operate at much higher frequencies.

Take, for example, an equipment working at VHF, say with f_c=100MHz and a bandwidth of 100kHz, presenting a smooth curve that allows its parameterization with a frequency spacing of Δf=5kHz. The inverse DFT of this band has $\Delta t \approx$5ns and T=200μs, which results in a top limit of 40,000 parameters to be used in the time-domain simulation. This number is already huge, but if some low frequency phenomena is also to be considered in the model, say at base-band with cut-off frequency at 2kHz, and with a pattern that requires a frequency spacing of 20Hz to be properly represented, then the total system impulse response would have T=50ms, keeping the previous sampling period Δt=5ns. The number of time domain parameters is then 10,000,000 (250 times more)! Fig. 5 illustrates this through an example of a system that clearly has slow and fast time scale characteristics.

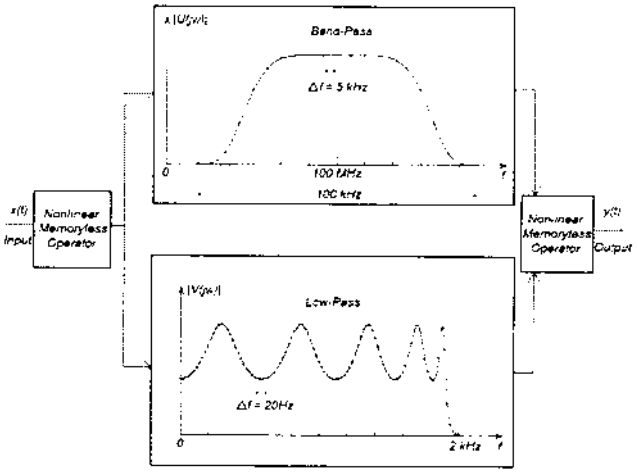

Fig. 5 - Example of a multi-rate system.

For a general multi-rate RF linear system, the number of time domain parameters is given by (3). It is equivalent to represent the full frequency span (from $-f_{máx}$ to $f_{máx}$) with the minimum frequency spacing Δf_{min}.

$$Number\ of\ Parameters = \frac{T_{max}}{\Delta t_{min}} = \frac{2 f_{max}}{\Delta f_{min}} \qquad (3)$$

Let us now consider the nonlinear case, restricting the nonlinearities up to the third-order kernel. The system model is given by equation (4).

$$\begin{aligned} y(t) = & \sum_{\tau=0}^{N-1} h_1(\tau) x(t-\tau) + \\ & + \sum_{\tau_1=0}^{M-1} \sum_{\tau_2=0}^{Q-1} h_2(\tau_1,\tau_2)\, x(t-\tau_1)\, x(t-\tau_2) + \\ & + \sum_{\tau_3=0}^{P-1} \sum_{\tau_2=0}^{Q-1} \sum_{\tau_3=0}^{R-1} h_3(\tau_1,\tau_2,\tau_3)\, x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3) \end{aligned} \qquad (4)$$

It is evident that $h_1(\tau)$ is the impulse response of the linear approximation of the system, to which corresponds $H_1(j\omega)$ in the frequency domain, after applying the DFT.

The second-order kernel $h_2(\tau_1,\tau_2)$ has a 2D domain, and $h_3(\tau_1,\tau_2,\tau_3)$ a 3D domain. They also have the respective frequency-domain representations $H_2(j\omega_1,j\omega_2)$ and $H_3(j\omega_1,j\omega_2,j\omega_3)$ which result from applying the multidimensional DFT [27].

Looking into the properties of the multidimensional DFT, it is observed that it preserves the same characteristics of the DFT, except that it is extended to a higher dimension domain. So, the problem of having a huge number of time domain parameters remains, and it gets even worse due to the domain dimension – that is, if $h_2(.)$ has a memory span of N elements on both τ_1 and τ_2 axes, it contains N^2 parameters, and if M is the memory span of $h_3(.)$ then it has M^3 time domain parameters!

If (4) is used as a RF model of a multi-rate nonlinear system, then it is necessary to constrain de domain of each kernel so that it gets highly reduced. Otherwise, there will be such a huge number of non-zero parameters that:

- the parameter identification process will probably be an impossible task;
- the amount of computer memory required to store the model is huge;
- the computation time during simulation is tremendous, even in high performance computers.

III. RESTRICTING THE NUMBER OF MODEL PARAMETERS

Fig. 5 already shows an example of a model constraint: the frequency domain representation has two clusters of non-zero parameters; all others are null. But, to what correspond these clusters in the respective time domain representation?

Fig. 6 shows the inverse DFT of both base-band and RF clusters. Since the responses of both low-pass and band-pass filters are supposed to interact, $v(t)$ is sampled at the same sampling frequency as $u(t)$, so $v(t)$ is represented with redundancy (we are considering that the model topology of Fig. 5 is not known a priori). This means that $v(t)$ can be approximated by a step function as shown in Fig. 7. In other words, $v(t)$ can be represented by 200 distinct parameters where each one is then repeated 50,000 consecutive times. This reduces, undoubtedly, the parameter identification burden, but in terms of simulation it still requires the convolution of all the 10 Msamples with the filters input signal.

This simulation process can also be simplified by grouping the input signal in 200 moving sums, each of which is multiplied by the respective parameter of $v(t)$ to perform the convolution. At each time instant, each moving sum only needs to add a new sample and subtract the tail sample, to the moving sum result of the previous epoch.

But it is also necessary to extract and use in simulation the 40,000 parameters of $u(t)$! If $U(j\omega)$ can have a smoother pattern then Δf can be higher, reducing the number of parameters in $u(t)$.

This constraint procedure, with the creation of clusters in the frequency-domain representation, can be extended to the nonlinear case. It is simply required to create n-dimensional clusters in the $H_n(j\omega_1,\ldots,j\omega_n)$ frequency domain kernels. But, contrary to the linear case, this is not an intuitive procedure in practice since multi-frequency signals are now being considered.

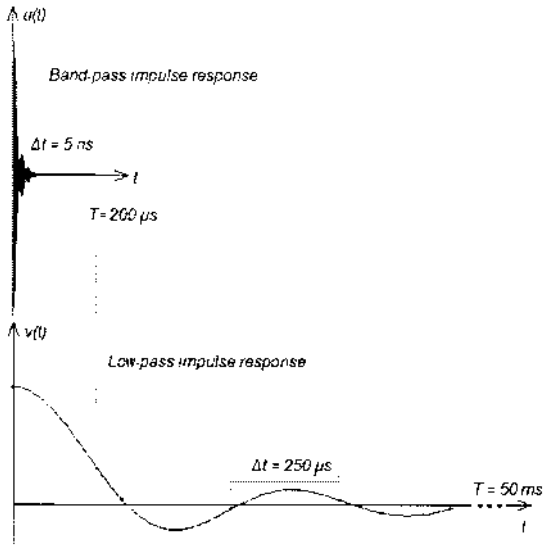

Fig. 6 - Impulse response of each band-pass and low-pass clusters.

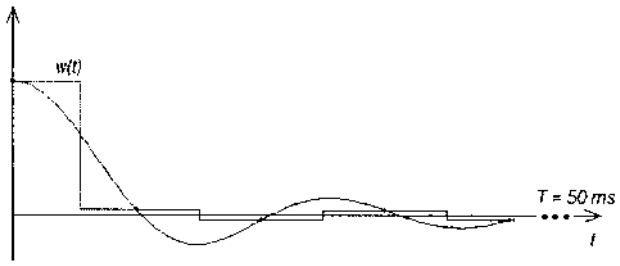

Fig. 7 - Stepwise approximation of v(t).

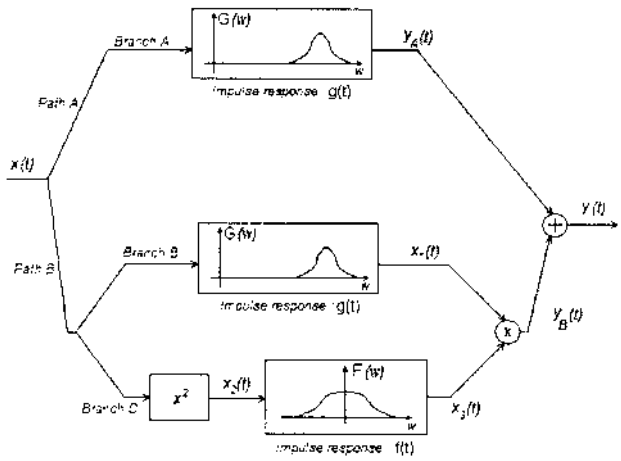

Fig. 8 - Example of a model topology.

Another way of restricting the information contained in a system is to use a model topology. Saying that a certain system fits the description of some model topology is, in

fact, an imposition on the information of the system that will be represented by that model. An example of what is called behavioural modelling with a priori knowledge of the system.

Let us consider an example of a model topology for a nonlinear system, given by Fig. 8.

This topology presents two main paths – path A models the linear behaviour of the system, which corresponds to the first-order kernel of (1); path B models the nonlinear characteristic of the system, which, in this case, is concentrated on the third-order kernel (its output is a double product of the input).

An immediate observation is that this topology restricts its usefulness to systems whose kernel orders other than first and third only have non-significant information. Nevertheless, it is known that several nonlinear interesting phenomena, in telecommunication systems and other areas, have most of their relevant information in those two kernels.

Notice that this topology can model multi-rate systems, as will be shown later, where low frequency components can influence the behaviour at in-band frequencies.

This model is then defined by:

- the model topology of Fig. 8;
- the N parameters of the band-pass linear filter of branches A and B – $g(t)$;
- the M parameters of the low-pass linear filter of branch B – $f(t)$.

So, given this topology, it should only be required $N+M$ parameters to model a system (and to simulate it).

Let us analyse the time-domain response of such topology, given a general input signal $x(t)$. Equation (5) shows the output of path A, and equation (6) the output of path B.

$$y_A(t) = \sum_{i=0}^{N-1} g(i)\, x(t-i) \tag{5}$$

$$y_B(t) = \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} g(i)\, f(j)\, x(t-i)\, x^2(t-j) \tag{6}$$

By matching the terms of equation (1) with equation (6) it becomes clear that (6) is the third-order kernel of a Volterra series expansion. Moreover, (7) gives a general definition of the third-order kernel that fits (6), where a restriction on the domain of the kernel is evident, as depicted in Fig. 9.

$$h_3(i,j,k) = \begin{cases} g(0)\cdot f(0) & , \ se\ i=j=k \\ \frac{1}{3}g(i)\cdot f(j) & , \ se\ i\neq j=k \\ \frac{1}{3}g(j)\cdot f(k) & , \ se\ j\neq k=i \\ \frac{1}{3}g(k)\cdot f(i) & , \ se\ k\neq i=j \\ 0 & , \ se\ i\neq j\neq k \end{cases} \tag{7}$$

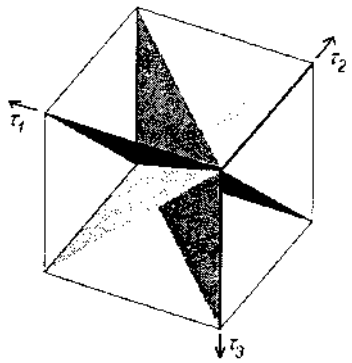

Fig. 9 - Domain of the h_3 time-domain kernel.

Notice that, for symmetry purposes, the term $g(i)f(j)$ was distributed evenly by the three admissible planes. Other distribution would also verify the matching with (6) as long as the sum of the three contributions would equal 1.

Assuming no particular specification to $g(t)$ and $f(t)$, it is obvious from (7) that the $h_3(\tau_1,\tau_2,\tau_3)$ kernel is formed by a bi-dimensional matrix that is copied into the three planes represented in Fig. 9. Taking, for example, $g(t)=u(t)$ of Fig. 6, and $f(t)=w(t)$ depicted in Fig. 7, this bi-dimensional matrix A, whose element (i,j) is given by $g(i)f(j)$, is plotted in Fig. 10. It is shown that each row equals the function $g(t)$ multiplied by a constant that changes only every 50,000 columns.

Similarly to what was exposed regarding the system of Fig. 5, the clustered information ($g(t)$ and $f(t)$ occupy two distinct and very separated frequency bands) is visible in the pattern of the matrix of Fig. 10. Again, it requires only $N+M$ parameters to be determined instead of $N\times M$ parameters (or even instead of the full cube of the $h_3(.)$ domain).

Continuing with the analysis in the time-domain, it is easy to observe that with the *multiple impulse input* identification method [8], the $N+M$ parameters of a system can be easily extracted without having to sweep the entire A matrix. In more general terms, if a certain model topology can be applied, then a specific and dedicated identification process can be used to extract the model parameters.

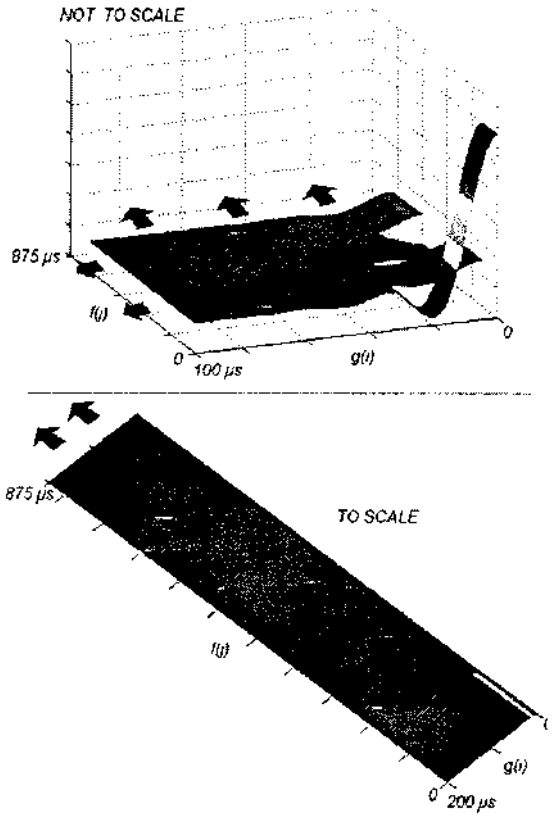

Fig. 10 - Bi-dimensional matrix A (only the upper envelope of $u(t)$ was considered).

In terms of simulation, two strategies are suggested to avoid making a large number of multiplications per epoch. The first is to implement an algorithmic scheme similar to the moving sum presented above. The second strategy, which looks more attractive, is to take advantage of the model topology. In fact, since the parameters of $g(t)$ and $f(t)$ can be identified separately, each branch of Fig. 8 can be evaluated independently, being the results of branch B and C multiplied and added to the result of branch A (moreover, the result of branch A is equal to the one of branch B).

Analysing, now, the topology of Fig. 8 in the frequency domain, it becomes again evident the restrictions that this topology imposes on the $H_3(.)$ kernel.

Using the *harmonic input method* [8], consider a three-tone input to path B, as shown in (8).

$$x(t) = e^{j\omega_1 t} + e^{j\omega_2 t} + e^{j\omega_3 t} \tag{8}$$

After proceeding with the math analysis of path B, and noticing that the final product is processed as a convolution in the frequency domain, expression (9) is reached. Then, the $H_3(j\omega_1, j\omega_2, j\omega_3)$ frequency-domain kernel is given by (10), according to the *harmonic input method.*

If $g(t)$ is a band-pass filter and $f(t)$ a low-pass filter, like those defined in Fig. 5, then $H_3(j\omega_1, j\omega_2, j\omega_3)$ has non-zero values on the coloured volumes depicted in Fig. 11. This evidently shows the restrictions that this topology imposes on the domain of $H_3(.)$.

$$\begin{aligned}
Y(\omega) = {} & G(\omega_1)\left(\begin{matrix} F(2\omega_1)e^{j3\omega_1} + F(2\omega_2)e^{j(\omega_1+2\omega_2)} + \\ + F(2\omega_3)e^{j(\omega_1+2\omega_3)} \end{matrix}\right) + \\
& + G(\omega_2)\left(\begin{matrix} F(2\omega_1)e^{j(2\omega_1+\omega_2)} + F(2\omega_2)e^{j3\omega_2} + \\ + F(2\omega_3)e^{j(\omega_2+2\omega_3)} \end{matrix}\right) + \\
& + G(\omega_3)\left(\begin{matrix} F(2\omega_1)e^{j(2\omega_1+\omega_3)} + F(2\omega_2)e^{j(2\omega_2+\omega_3)} + \\ + F(2\omega_3)e^{j3\omega_3} \end{matrix}\right) + \\
& + 2G(\omega_1)\left(\begin{matrix} F(\omega_1+\omega_2)e^{j(2\omega_1+\omega_2)} + F(\omega_1+\omega_3)e^{j(2\omega_1+\omega_3)} + \\ + F(\omega_2+\omega_3)e^{j(\omega_1+\omega_2+\omega_3)} \end{matrix}\right) + \\
& + 2G(\omega_2)\left(\begin{matrix} F(\omega_1+\omega_2)e^{j(\omega_1+2\omega_2)} + F(\omega_1+\omega_3)e^{j(\omega_1+\omega_2+\omega_3)} + \\ + F(\omega_2+\omega_3)e^{j(2\omega_2+\omega_3)} \end{matrix}\right) + \\
& + 2G(\omega_3)\left(\begin{matrix} F(\omega_1+\omega_2)e^{j(\omega_1+\omega_2+\omega_3)} + F(\omega_1+\omega_3)e^{j(\omega_1+2\omega_3)} + \\ + F(\omega_2+\omega_3)e^{j(\omega_2+2\omega_3)} \end{matrix}\right)
\end{aligned} \tag{9}$$

$$\begin{aligned}
& H_3(\omega_1,\omega_2,\omega_3) = \\
& = \frac{1}{3}\left(G(\omega_1)F(\omega_2+\omega_3) + G(\omega_2)F(\omega_1+\omega_3) + G(\omega_3)F(\omega_1+\omega_2)\right)
\end{aligned} \tag{10}$$

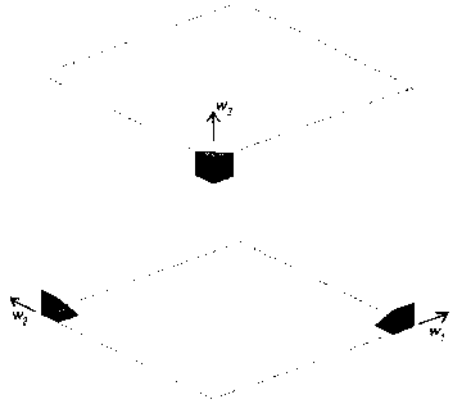

Fig. 11 - Domain of the H3 frequency-domain kernel.

VI. CONCLUSIONS

When multi-rate phenomena are being studied in a nonlinear (or even on a linear) system, some care must be taken when an RF model of the system is considered and implemented by means of a nonlinear FIR filter. As shown in this paper, the number of model parameters easily grows to incredible values that obviate any system identification procedure, and deeply compromises the computation time required to simulate the system behaviour.

To prevent that, some strategy must be used to strongly restrict the domain of the nonlinear FIR filter kernels. A suggested approach is to consider a specific a priori topology for the model. The model topology is, by itself, a

specification of how the system relevant information is (or should be) distributed. A proper model topology can largely reduce de number of parameters necessary to model the system, and also to significantly reduce the epoch-to-epoch simulation time. Dedicated parameter extraction procedures can also be specified to ease the identification process required to determine the parameter values that best model a certain system (according to that topology).

However, to guarantee the desired model predictive capabilities, when using a determined topology it is necessary to verify that it can, in fact, represent the system behaviour.

REFERENCES

[1] V. Rizzoli and A. Neri, "State of the Art and Present Trends in Nonlinear Microwave CAD Techniques", IEEE Trans. on Microwave Theory and Tech., Vol. MTT-36, No. 2, pp.343-365, 1988.

[2] M. S. Nakhla and J. Vlach, "A Piecewise Harmonic Balance Technique for Determination of Periodic Response of Nonlinear Systems", IEEE Trans. on Circuits and Systems, Vol. CAS-23, No. 2, pp.85-91, 1976.

[3] V. Rizzoli, A. Neri and F. Mastri, "A Modulation-Oriented Piecewise Harmonic Balance Technique Suitable for Transient Analysis and Digitally Modulated Analysis", 26th European Microwave Conference Proc., pp.546-550, Prague, 1996.

[4] J. C. Pedro and N. B. Carvalho, "Simulation of RF Circuits Driven by Modulated Signals Without Bandwidth Constraints", 2002 IEEE International Microwave Symposium Dig., Seattle, 2002.

[5] D. Sharrit, "New Method of Analysis of Communication Systems", IEEE MTT-S Nonlinear CAD Workshop, 1996.

[6] L. Chua, "Nonlinear Circuits", IEEE Trans. on Circuits and Systems, Vol. CAS-31, No. 1, pp.69-87, 1984.

[7] V. Mathews and G. Sicuranza, *Polynomial Signal Processing*, John Wiley & Sons, Inc., New York, 2000.

[8] J. C. Pedro and N. B. Carvalho, Intermodulation Distortion in Microwave and Wireless Circuits, Artech House, Norwood, 2003.

[9] T. J. Aprille and T. N. Trick, "Steady-State Analysis of Nonlinear Circuits With Periodic Inputs", Proceedings of the IEEE, Vol. 60, No.1, pp.108-114, 1972.

[10] S. A. Maas, Nonlinear Microwave Circuits, Artech House, Norwood, MA, 1988.

[11] M. Schetzen, The Volterra and Wiener Theories of Nonlinear Systems, John Wiley & Sons, New York, 1980.

[12] N. B. Carvalho and J. C. Pedro, "Multi-tone Frequency Domain Simulation of Nonlinear Circuits in Large and Small Signal Regimes", IEEE Trans. on Microwave Theory and Tech., Vol. MTT-46, No. 12, pp.2016-2024, 1998.

[13] C. R. Chang and M. B. Steer, "Frequency-Domain Nonlinear Microwave Circuit simulation Using the Arithmetic Operator Method", IEEE Trans. on Microwave Theory and Tech., Vol. MTT-38, No. 8, pp.1139-1143, 1990.

[14] V. Rizzoli, F. Mastri, E. Furini and A. Costanzo, "A Krylov-Subspace Technique For The Global Stability Analysis of Large Nonlinear Microwave Circuits", 2001 IEEE International Microwave Symposium Dig., pp.435-438, 2001.

[15] N. B. Carvalho and J. C. Pedro, "Analysis and Measurement of Multi-tone Intermodulation Distortion of Microwave Frequency Converters", 2001 IEEE International Microwave Symposium Dig., pp. 1671-1674, 2001.

[16] D. Hente and R. H. Jansen, "Frequency Domain Continuation Method for the Analysis and Stability Investigation of Nonlinear Microwave Circuits", IEE Proceedings-H Microwaves Antennas and Propagation, Vol. 133, No. 5, pp. 351-362, 1986.

[17] P. J. Rodrigues, Computer Aided Analysis of Nonlinear Microwave Circuits, Artech House, Inc., Norwood, MA, 1998.

[18] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, Norwood, 2000.

[19] G. Cybenko, "Approximation by Superpositions of a Sigmoidal Function", *Math. Control Signals Systems*, vol. 2, pp.303-314, 1989.

[20] K. Hornik, M. Stinchcombe and H. White, "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks*, vol. 2, pp.359-366, 1989.

[21] Y. Fang, M. C. Yagoub, F. Wang and Q. J. Zhang, "A New Macromodeling Approach for Nonlinear Microwave Circuits Based on Recurrent Neural Networks", *IEEE Trans. on Microwave Theory and Tech.*, vol. MTT-48, pp.2335-2344, Dec. 2000.

[22] D. Schreurs, N. Tufillaro, J. Wood, D. Usikov, L. Barford and D. E. Root, "Development of Time Domain Behavioural Non-Linear Models for Microwave Devices and ICs from Vectorial Large-Signal Measurements and Simulations", *Europ. Gallium Arsenide and other Semiconductors Applications Symp. Dig.*, pp.236-239, Oct. 2000.

[23] V. Rizzoli, A. Neri, D. Masotti and A. Lipparini, "A New Family of Neural Network-Based Bidirectional and Dispersive Behavioral Models for Nonlinear RF/Microwave Subsystems", *Int. Jour. of RF and Microwave CAE*, vol. 12, pp.51-70, 2002.

[24] J. Xu, M. Yagoub, R. Ding and Q. J. Zhang, "Neural-Based Dynamic Modeling of Nonlinear Microwave Circuits", *IEEE Trans. on Microwave Theory and Tech.*, vol. MTT-50, pp.2769-2780, Dec. 2002.

[25] F. H. Harris, Multirate Signal Processing for Communication Systems, Prentice Hall PTR, 2004.

[26] A. Oppenheim and R. Schafer, Discrete-Time Signal Processing, Prentice Hall, Englewood Cliffs, 1999.

[27] R. Tolimieri, M. An and C. Lu, Mathematics of Multidimensional Fourier Transform Algorithms, Springer Verlag, New York, 1997.